University of Bath

PHD

Innovative Microelectronic Signal Processing Techniques for the Recording and Analysis of the Human Electroneurogram

Metcalfe, Benjamin

Award date:
2016

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 05. May. 2023

Innovative Microelectronic Signal Processing Techniques for the Recording and Analysis of the Human Electroneurogram

submitted by

Benjamin William Metcalfe

for the degree of Doctor of Philosophy

University of Bath

Department of Electronic and Electrical Engineering

January 2016

COPYRIGHT

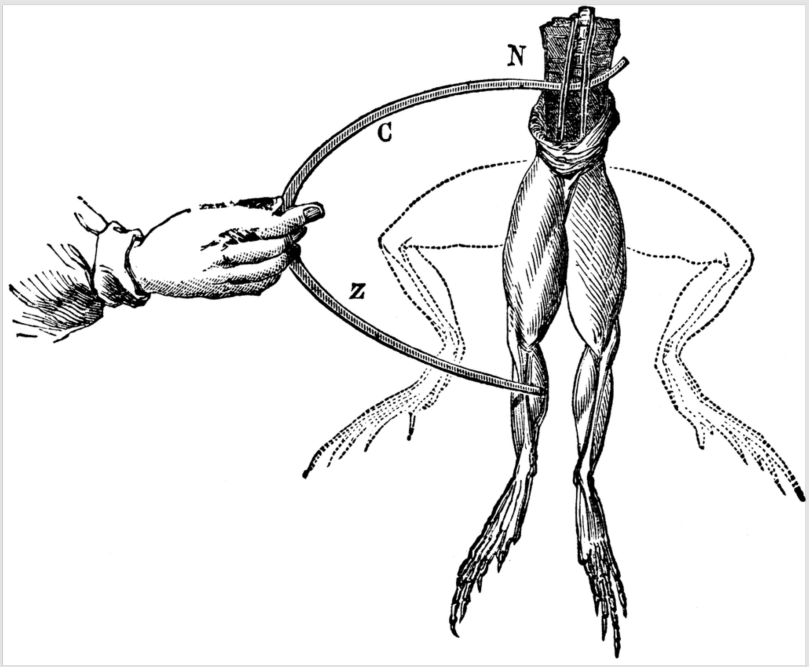
Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

Signature of Author .....

Benjamin William Metcalfe

This thesis is dedicated in loving memory of my father
SIMON JOHN METCALFE 1969 - 2011



I am attacked by two very opposite sects -
the scientists and the know-nothings.
Both laugh at me - calling me "the frogs' dancing-master".
Yet I know that I have discovered one of the greatest forces in nature.
LUIGI GALVANI - 1792

Acknowledgements

This work would not have been possible without the help and support of many people over the past three years.

Firstly, I would like to thank my doctoral supervisors Dr. Chris Clarke and Prof. John Taylor, under whose guidance and supervision this work has been completed. They have provided me with continued support and valuable comments, and have spent a good deal of time correcting my grammar. It has been a pleasure to work alongside them.

I must also thank Prof. Nick Donaldson from University College London who has often provided a guiding hand, and in suffering through the triumphs and the failures with me, has been like another supervisor.

I would also like to thank Dr. Daniel Chew from the University of Cambridge for his invaluable surgical skills and for his patience in the operating room when making our recordings from *rat*. I am also grateful for the assistance of Dr. Thomas Nielsen at the University of Aalborg for arranging the experiments performed with *pig* and for Dr. Martin Schuettlers support in providing me with his data recorded from *frog*.

I am profoundly fortunate to have been funded through my doctoral study by Eur. Ing. Dr. Brian Nicholson, who has provided me with unwavering and unquestioning support for the last three years and has asked for very little in return. I am also fortunate to have been supported by Mr. Roger Whorrod OBE and Mrs. Sue Whorrod, whose philanthropy enabled my undergraduate studies at the University of Bath. I hope that this thesis and the work contained within it is a worthy testament to this support; I am forever in their debt.

Of course I must also thank my family, friends and fellow students for their support and advice. I must thank my grandfather, Bill Honey, whose garden shed full of old radio valves first got me interested in electronics. I am especially grateful to my partner, Elli Jefferson, who has supported me unquestionably during the writing of this thesis and my mother, Sarah Parr, who always encouraged me to follow my dreams and inspired me to achieve all that I have.

Abstract

Injuries involving the nervous system are among the most devastating and life altering of all neurological disorders. The resulting loss of sensation and voluntary muscle control represent a drastic change in the individuals lifestyle and independence. Spinal cord injury affects over two hundred thousand people within the United States alone. While there have been many attempts to develop neural interfaces that can be used as part of a prosthetic device to improve the quality of life of such patients and contribute to the reduction of ongoing health care costs, the design of such a device has proved elusive. Direct access to the spinal cord requires potentially life threatening surgery during which the *dura*, the protective covering surrounding the cord, must be opened with a resulting high risk of infection. For this reason research has been focussed on the stimulation of and recording from the peripheral nerves in an attempt to restore the functionality that has been lost through spinal cord injury.

This thesis is concerned with the current status and limitations of peripheral nerve interfaces that are designed for recording electrical signals directly from the nervous system using a technique called *velocity selective recording*. This technique exploits the relationship between axonal diameter, which is linked via anatomy to function, and the speed with which the axon conducts excitation. New techniques are developed that improve current methods for identifying and simulating neural signals and power efficient implementations of these methods are presented in modern microelectronic platforms. Results are presented from pioneering experiments in *rat* and *pig* that for the first time demonstrate the recording and analysis of the physiological electroneurogram using velocity based methods. New methods are developed that enable the extraction of neuronal firing rates and thus the extraction of the information encoded within the nervous system.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC Analogue to Digital Converter

AP Action Potential

ASIC Application Specific Integrated Circuit

AWGN Additative White Gaussian Noise

CAP Compund Action Potential

CMOS Complimentary Metal Oxide Semiconductor

CPLD Complex Programmable Logic Device

CSN Correlated Stochastic Noise

DAC Digital to Analogue Converter

ECG Electrocardiogram

EEG Electroencephalogram

EMG Electromyogram

ENG Electroneurogram

ENOB Effective Number of Bits

FES Functional Electrical Stimulation

FINE Flat Interface Nerve Electrode

FIR Finite Impulse Response

FPGA Field Programmable Gate Array

IC Integrated Circuit

IIR Infinite Impulse Response

IVS Intrinsic Velocity Spectrum

JTAG Joint Test Action Group

LIFE Longitudinally Implanted Intrafascicular Electrode

MEC Multiple Electrode Cuff

NEO Nonlinear Energy Operator

OU Ornstein-Uhlenbeck

PCA Principal Component Analysis

PDMS Polydimethylsiloxane

PWM Pulse Width Modulation

RAM Random Access Memory

RMS Root Mean Square

RTL Register Transfer Level

SNR Signal-To-Noise-Ratio

SPINE Slowly Penetrating Interfascicular Electrode

TDNN Time Delay Neural Network

TMAP Transmembrane Action Potential

TIME Transverse Intrafascicular Multichannel Electrode

TTL Transistor-Transistor Logic

USB Universal Serial Bus

VLSI Very Large Scale Integrated circuit

VSR Velocity Selective Recording

WRE Width Reduction Envelope

Chapter 1

Introduction

1.1 Motivation for this thesis

The existence of the human nervous system has been recognised since *c.* 500 BC when the Greek philosopher Alcmaeon made the earliest recorded descriptions of the optic nerve. Since this time mankind has expended a great deal of effort exploring the mechanisms by which the various connections of axons and dendrites manifest themselves into life. For many years the study of the nervous system was solely through dissections, observed with the unaided eye; however during the XVII and XVIII centuries advances in the fields of optics and electrical phenomena were to revolutionise neural investigations. In 1791 Luigi Galvani published his seminal work entitled *De viribus electricitatis in motu musculari commentarius* in which he detailed experiments observing that the legs of a *frog* could be made to spasm by the application of electricity [1]. After dedicating much of his life to the study of this *animal electricity*, Galvani made the important discovery that electricity is not only able to control the nervous system but it is in fact the very mechanism by which it functions. Combined with the work of Volta, with whom Galvani worked closely, the field of *electrophysiology* was born.

Towards the end of the XIX century it became clear not only that nerves and muscles could be activated by electrical energy but that their normal activity was always accompanied by changes in electric potential [2]. It was not until the 1930s that it was conclusively shown by Alan Hodgkin and Andrew Huxley that electrical conduction was the direct method by which nerves convey information around the body [3]. Further observation demonstrated that by connecting nerves together it becomes possible to process information, form decisions and take actions. It is these fundamental processes that when interconnected give rise to the enormous, and perhaps unfathomable, complexity of the mind. There is then a great interest in the study and understanding of nerves so that we might better be able to treat their ailments more safely as well as developing an intrinsic understanding of life systems.

The electrical signals carried through the body can be recorded in a number of differ-

ent ways, depending on the diagnostic requirements of the study. Broadly speaking the recording of the electrical signals may be broken down in three main categories based on their origins [4],

Electromyography (EMG) - *Electromyography* refers to the electrical signals that are responsible for the contraction or flexion of muscle fibres. For example the Electrocardiogram (ECG) which is often recorded from electrodes placed on the dermis, records the contraction of the heart muscles.

Electroencephalography (EEG) - *Electroencephalography* is the recording of electrical activity from the brain often using surface electrodes similar to those employed in the recording of the EMG. The recording of the EEG is a vital diagnostic tool for the analysis of neurological disorders such as epilepsy and Alzheimers disease.

Electroneurography (ENG) - A relatively recent endeavour (at least in so far as *in vivo* experimentation), *electroneurography* is the recording of the electrical activity directly from a group or bundle of nerves. Generally the ENG is recorded from either the *cauda equina* or the peripheral nervous system.

Injuries to the nervous system can cause permanent loss of sensation and motor function, placing them amongst the most devastating of all injuries. The effects are often irreversible and may leave the victim with permanent life altering paralysis.

Over two hundred thousand people within the United States of America alone are affected by spinal cord injury and the total direct costs associated with lifetime care can exceed seven billion dollars each year [5]. In focal spinal cord injuries there is often dramatic paralysis of the legs, abdomen and bowels; yet the peripheral nerves connected to these areas are commonly still intact. There are numerous approaches aimed at reducing the paralyzing effect of injury and promoting regrowth of nerve fibres, however inhibitory factors such as inflammation and glial scarring limit the effectiveness of post injury neuroregeneration. In the absence of neuroregeneration, *functional electrical stimulation* (FES) aims for the full or partial restoration of lost function by stimulation of the intact nerves below the lesion of the spinal cord with electrical impulses.

There have been many attempts to apply FES to the development of neural interfaces that can be used as part of a prosthetic device to improve quality of life, however such devices have proven to be difficult to manufacture due to the lack of a reliable recording mechanism that would allow for feedback within the system. The most straightforward way to interface with the nervous system is via direct access to the spinal cord, which requires potentially life threatening surgery, during which the *dura* (the protective covering surrounding the cord) must be removed, presenting a high risk of infection. An alternative method that circumvents this invasive surgery is to stimulate and record from the peripheral nervous system near the site of the target organ or system, in an attempt to restore the functionality that has been lost through spinal cord injury.

Direct electrical stimulation of the peripheral nerves has seen success in a number of applications, predominantly in motor neuroprostheses such as the types used for foot-drop correction [6] or the restoration of hand grasp in tetraplegic patients [7]. However, at present the effectiveness of such prostheses remains limited. There is no mechanism by which neural signals may be recorded and interpreted, thereby allowing a closed loop system where, for example, the contact force exerted by the hand can be measured. If a suitable recording mechanism could be devised, then it would vastly improve the functionality of current neuroprosthetic devices. Of course the ability to record from the peripheral nerves also holds great importance in the study and understanding of nervous system disorders and diseases. The nervous system, including the brain, is often described as one of the final frontiers of modern science.

1.2 Contributions of this thesis

The following can be identified as major contributions of this thesis to the advancement of neural signal recording:

1. Recordings of the physiological ENG were made *in vivo* from *rat* using multiple hook electrodes connected in isolated dipoles and immersed in oil. These recordings show individually identifiable action potentials that were analysed automatically using newly developed signal processing methods. Modulation of the ENG is elicited via direct stimulation of the *derma* of the animal and statistically significant changes in the recorded ENG are shown to correlate with the stimulation event. These recordings represent the first application of velocity selective recording to physiological ENG.

2. Pioneering recordings of both electrically evoked and physiological ENG have been made from the intact right vagus nerve of *pig*. These recordings were made with a chronically implantable electrode structure (a cuff electrode) and extend the work presented in *rat* to data obtained with an implantable electrode structure.

3. A new method is described that enables the firing rates of individual neurons to be extracted in real time using a velocity based signal processor. Currently available methods for analysing the velocity content of neural recordings provide no method by which firing rates may be extracted. The new method extends the current methods to provide this functionality and has been validated using physiological recordings of ENG made in *rat*.

4. A novel discrete time filter has been developed that provides an efficient method for finding the *centre of gravity of pulses*. This filter is demonstrated as a promising method for the temporal alignment of action potentials that is substantially more resilient to both correlated and uncorrelated noise than existing methods. A multiplier free implementation is described and detailed power and resource measurements are

presented for a typical implementation.

5. Velocity selective recording is compared and contrasted with traditional morphological spike sorting techniques. Both methods are applied to simulated data as well as to data recorded *ex vivo* (from explanted neurons) from *frog*, and it is shown that both methods classify neural activity based on axonal diameter and may be equally well applied to the same data sets. The data recorded from *frog* were recorded in another study but have been used here for comparative purposes with permission from the original authors.

6. A high speed hardware based simulation system has been developed that extends existing technology using modern microelectronic systems to provide a high level of configurability. The simulation system makes use of recent recordings made in *rat* to produce realistic test patterns for the design of velocity based signal processors. The simulator addresses a number of key issues with existing designs and aims to reduce the reliance on testing with animal models.

A list of publications inspired by the work of this thesis is given below:

- B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. de N. Donaldson, and J. T. Taylor, "A new method for spike extraction using velocity selective recording demonstrated with physiological ENG in Rat.," *J. Neurosci. Methods*, vol. 251, pp. 47-55, Aug. 2015.

- B. Metcalfe, D. Chew, C. Clarke, N. Donaldson, and J. Taylor, "An enhancement to velocity selective discrimination of neural recordings: Extraction of neuronal firing rates.," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2014, pp. 4111-4114, Aug. 2014.

- B. Metcalfe, D. Chew, C. Clarke, N. Donaldson, and J. Taylor, "Fibre-selective discrimination of physiological ENG using velocity selective recording: Report on pilot rat experiments.," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 2645-2648. 2014.

- J. Taylor, B. Metcalfe, C. Clarke, D. Chew, T. Nielsen, and N. Donaldson, "A Summary of Current and New Methods in Velocity Selective Recording (VSR) of the Electroneurogram (ENG).," *2015 IEEE Int. Symp. VLSI*, 2015.

Pending Publications:

- B. W. Metcalfe, C. T. Clarke, N. de N. Donaldson, and J. T. Taylor, "Improved Method for Neural Spike Alignment: The Centroid Filter" *IEEE Trans Neural Systems & Rehabilitation Engineering.*

- B. W. Metcalfe, T. Neilsen, J. Struijk, C. T. Clarke, N. de N. Donaldson, and J. T. Taylor, "Report on Experiments to Record from the Peripheral Nerves of Pig in vivo." *IOP Journal of Neural Engineering.*

References

[1] L. Galvani, *De Viribus Electricitatis In Motu Musculari Comentarius*. Apud Societatem Typographicam, 1791.

[2] J. Nicholls, A. Martin, and B. Wallace, *From Neuron To Brain*. Sinauer Associates, third ed., 1998.

[3] A. Hodgkin and A. Huxley, "Action potentials recorded from inside a nerve fibre," *Nature*, vol. 144, pp. 710–711, 1939.

[4] R. Carpenter, *Neurophysiology*. Arnold Publishers, fourth ed., 2003.

[5] C. Sadowsky, J. W. McDonald, O. Volshteyn, and L. Schultz, "Spinal cord injury.," *Disability and rehabilitation*, vol. 24, pp. 680–7, Sept. 2002.

[6] M. Hansen, M. K. Haugland, and F. Sepulveda, "Feasibility of using peroneal nerve recordings for deriving stimulation timing in a foot drop correction system.," *Neuromodulation : journal of the International Neuromodulation Society*, vol. 6, pp. 68–77, Jan. 2003.

[7] M. Haugland, A. Lickel, R. Riso, M. M. Adamczyk, M. Keith, I. L. Jensen, J. Haase, and T. Sinkjaer, "Restoration of lateral hand grasp using natural sensors.," *Artificial organs*, vol. 21, pp. 250–3, Mar. 1997.

Chapter 2

The Human Nervous System

2.1 Overview - Systema Nervosum

In order to discuss and assess appropriately the different techniques for neural recordings it is essential to have a good understanding of the basic structure and operation of the nervous system (*systema nervosum*). This chapter provides a brief overview of the nervous system and of existing techniques for recording from the peripheral nervous system. Although it is essentially continuous the nervous system can be divided into parts, regions and systems.

The *encephalon* or brain and *medulla spinalis* or spinal cord form the central nervous system (*systema nervosum centrale*) extending from which are twelve cranial nerves and thirty one spinal nerves that form the peripheral nervous system (*systema nervosum periphericum*). The division of the nervous system in this way is largely functional; the peripheral nervous system consists of relatively simple conductors connecting peripheral receptor and effector organs to each other through the intermediation of the brain and spinal cord. The peripheral nervous system includes the *craniospinal* and *autonomic* nerves and their associated ganglia, together with their connective tissue sheaths. All lie peripheral to the *pia*l covering of the central nervous system [1].

The peripheral nervous system can be further classified as containing both the somatic and the autonomic nervous system. The somatic nervous system is responsible for voluntary body movements via the skeletal muscles; it includes the nerves responsible for muscle contraction as well as the motor neurons contained within the skin. The autonomic nervous system is responsible for the involuntary bodily functions such as physiologic respiration, cardiac function and mucus production. Not all functions of the autonomic nervous system are involuntary however - take for example the process of sexual arousal. The primary physical effects of arousal such as vasodilation and an increase in cardiac output are an autonomic response to a conscious and voluntary thought [2]. The autonomic nervous system can be further divided into the *somatosensory* system, which is of interest in the study of neuroprostheses, as well as the *sympathetic* and the *parasympathetic* systems. It

is the somatosensory system that contains the primary sensory and motor neurons that are responsible for sensation and movement.

Nerves are often classified by the direction in which they propagate information; motor neurons are classified as *efferent* or "exiting" from the brain and sensory neurons are *afferent* or "entering" the brain. In some simple organisms such as the *hydra*, the sensory neurons may connect directly to the motor neurons with very limited processing. Such connections are responsible for reflex reactions in many animals [3]. The nervous system of *man* by comparison features a highly complex system of a great many interconnected neurons - over 86 billion in total - that facilitate the detailed decision making processes and other higher functions that we associate with the conscious mind [4].

The nervous system is both mechanically and chemically fragile and must be kept in a tightly controlled environment. The most sensitive parts of the central nervous system are encased in the protective structures of the *cranium* and the *spinal canal* of the vertebral column; whereas the peripheral nerves are most often encased in a thin protective *epineurium* and are surrounded by soft tissues. There are two main causes of chronic failure within the nervous system. The most common cause is severe mechanical trauma to the medulla spinalis, which can result in crush damage to individual nerves and may lead to *Wallerian degeneration* in which the part of the axon that has been crushed degenerates distally to the injury leading to the eventual death of the nerve [5]. If the injury occurs in an area associated with cardiac function or respiration, death is almost instantaneous. If however the injury occurs lower down the medulla spinalis or even within the peripheral nervous system, the most likely outcome is localised or total paralysis. This paralysis is seldom reversible and can become a significant factor in both ongoing health care costs and the associated quality of life perceived by the injured. Notably the sections of the nerve either side of the injury often remain fully functional and the chronic prognosis for these nerves is good [6].

The second most common cause of chronic failure is neurodegeneration brought about by diseases such as *Multiple Sclerosis* in which the nerves are stripped of their insulating layer (*demyelinated*) that is required for proper function [7]. There is no cure for such diseases and most treatments focus on palliative care and the reduction of further damage.

In recent years the development of neural interfaces has facilitated the design of neural prosthesis that aim to increase the quality of life for patients that have suffered some form of injury or degeneration of the nervous system. A closed loop neural prosthesis that both records from and stimulates the nervous system has the capability to partially or fully restore functionality lost from this type degeneration.

2.2 The Neuron

Neurons from different parts of the nervous system vary considerably in morphology, but they possess many common features. They all have a *soma* or cell body that is relatively large and contains a nucleus surrounded by a cytoplasmic mass. Extending from the soma are fine filaments that are collectively termed *neurites*, the first of these are the *dendrites*, which are multiple in number and conduct excitation towards the soma and so act as the inputs to the neuron. The second type of neurite is the *axon*, which is singular and conducts excitation away from the soma.

The axons usually end in several fine *telodendria* that are coincident with the dendrites of other neurons; forming *synapses* with either effectors such as muscles fibres or with glands. Neurons range in size; with the smallest somata at 7 μm in diameter and the largest cells of the spinal ganglia at over 120 μm, whilst also varying in length from a few microns to a few tens of centimetres [1].

The axon, which is the longest part of the neuron, is designed for carrying excitation over a long distance and at high speed in a manner that is unachievable using conventional chemical or hormonal diffusion. Information is transmitted along the axon in the form of *action potentials* (APs), small impulses that propagate along the axon from the soma to the telodendria [3]. Larger axons are covered in an insulating sheath of *myelin*, a lipid substance that improves the conduction of APs by acting as an electrical insulator [8], this myelin sheath is interrupted at regular intervals called the nodes of *Ranvier*.

The transmission of excitation along the length of the axon is an electro-chemical process based on the exchange of charged ions across cell membranes, in this way it produces a longitudinal flow of current and so may be modelled to a first approximation as electrical conduction. When the AP reaches the telodendria, a small amount of chemical transmitter is released, which can be detected by the coincident dendrites of other neurons which can in turn generate further APs in their own axons. At either end of this process chemical signals are used to communicate with cells of other types such as muscles or cutaneous touch receptors. Each axon is enclosed within a *fascicle* (or bundle) with other axons that is contained within the protective internal epineurium. The fascicles are in turn bundled together along with blood vessels into the nerve and further enclosed within the external epineurium.

2.3 Neurons as Conductors of Electricity

In order to understand how an effective neural recording system may be constructed it is necessary to understand the basic processes by which the neurons transmit information along their length. The biochemical processes can be approximated and modelled using elements from traditional circuit theory, this can be done by realising that the charged

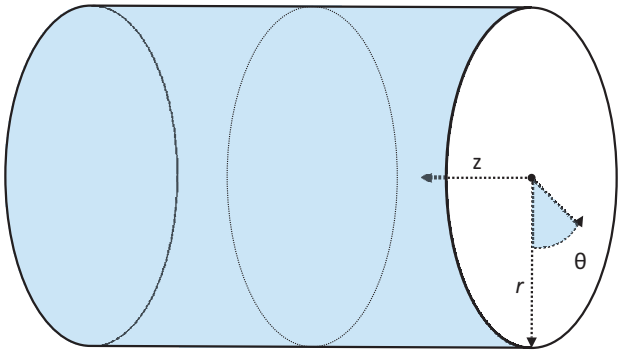


Figure 2.3.1: Cylindrical coordinate system suitable for analysing longitudinal current.

ions present in different concentrations on either side of the cell membrane form analogies to voltages and currents. In this chapter the mechanisms behind AP propagation are described in terms of basic electrical models and equivalent circuit parameters are provided for the purposes of quantitative analysis.

2.3.1 Passive Conduction

The transmission of excitation along the axon is often modelled using the *core-conductor* model. The core-conductor model refers broadly to a set of concepts used for analysing the longitudinal spread of electrical current along the length of the axon. If an electrical model of the axon is to be described based on concepts from basic electrical circuits then a few reasonable assumptions must be made. Firstly it is assumed that the axon features axial symmetry, that is, $\partial/\partial\Theta = 0$ (where Θ is the azimuth angle, see Figure 2.3.1). This assumption constrains the field quantities to be functions purely of r and z (i.e cylindrical coordinates). Another assumption is that these field quantities are independent of r and in fact vary only as a function of the axial coordinate z.

The axon is an imperfect conductor, and so the spreading of current over a fixed distance results in an attenuation. In the axon this attenuation is dictated by a number of factors that can be derived from the diameter and membrane properties of the axon. For a large and well insulated axon, an electrical impulse will cover a long distance before succumbing to appreciable attenuation; conversely a poorly insulated axon with a small diameter will attenuate an electrical impulse over a much shorter distance. There is a capacitance between the axon and the extracellular fluid that surrounds it, which will further affect the transmission of electrical impulses.

The axons of most of the nerves within the peripheral nervous system are covered by a high resistance and low capacitance insulator, formed by a thin layer of the lipid substance called *myelin* [3, 8, 10]. The myelin sheath forms a very effective insulator from the surrounding tissue and constrains the propagation of excitation to within the axon. This sheath is not continuous; it is interrupted at regular intervals by the nodes of Ranvier that provide a return path for current through the extracellular fluid. The insulating effect of the myelin

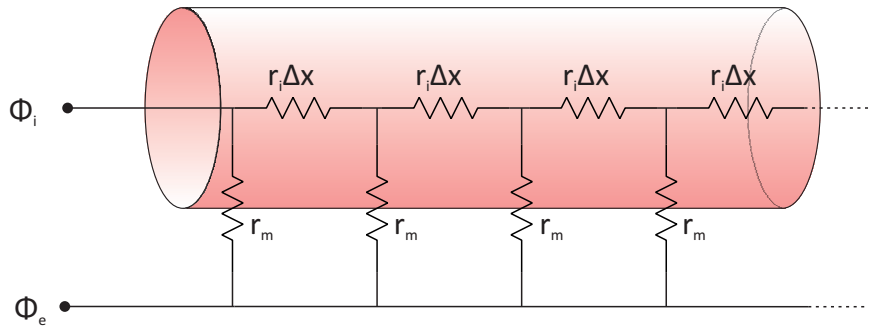


Figure 2.3.2: Simplified lumped parameter model of an axon.

sheath is also required for rapid conduction of excitation along the nerve, thus myelinated axons tend to occur in pathways of the nervous system where either speed or conduction distance are important.

A numerical model of the passively conducting axon can be made by drawing on the analogy between axons and electrical conductors. This analogy was first demonstrated by Hodgkin and Rushton who observed the longitudinal spread of applied current along axons from *Nephropidae* [11]. In a traditional cable, for example the type often used in underwater telegraphy, there is a core conductor formed from copper that is surrounded by an insulating plastic. A further conductor surrounds this insulator and acts as the return path for the flow of electrical current. Quantitatively however, there are numerous key differences between the axon and the telegraphic cable. In a cable with a core conductor made of a very highly conductive material, the insulator may be a material with a very high impedance. The capacitance of the insulator can be made very low by increasing the thickness of its material; as a result an injected current will travel a great distance along such a cable, since the resistance to longitudinal current flow is low and very little current is lost due to leakage through the insulator.

In the axon, the core conductor, or *axoplasm*, is a saline solution with a conductivity much lower than that of copper, whilst the tissue surrounding the outside of the axon has a much higher conductivity. The membrane, for unmyelinated axons, is a poor insulator and is relatively thin and so has appreciable capacitance. Using this simple model, a current injected into one end of an axon will fail to travel very far due to the low conductance of the axoplasm, as well as a progressive loss of current through the insulating membrane. To provide a quantitative comparison the resistivity of copper is 1.68×10^{-2} Ωm^{-1} and the longitudinal resistivity of a typical large axon is 6000 $\text{M}\Omega\text{m}^{-1}$ [3].

In the first instance we can define two resistance elements that describe the axon under passive conduction: the longitudinal resistance of the axoplasm r_i (or the intracellular resistance) and the transverse resistance of the membrane r_m as shown in Figure 2.3.2. In terms of the cylindrical volume conductor it is assumed that r_i is independent of Θ and r and that r_m is independent of Θ and z. Making the assumption that the extracellular fluid surrounding the axon has negligible impedance to the flow of current, the outer ends of

Table 2.3.1: Typical voltage attenuation along a myelinated axon that is uninterrupted by nodes of Ranvier and assuming purely passive conduction.

Distance (mm)	Amplitude (Volts)	Attenuation (%)
0	1	0
4	0.368	63.2
8	0.135	86.5
16	0.018	98.2

the individual transmembrane resistances may be shorted together. The model shown in Figure 2.3.2 resembles the traditional lumped model that is often applied to the analysis of transmission lines. If a potential V_0 is applied between Φ_i (the intracellular potential) and Φ_e (the extracellular potential) at one end of the model, then at each section the current generated by the applied potential will flow through both the longitudinal intracellular resistance and through the transverse transmembrane resistance. The ratio of these two resistances will define the proportion of current that flows into each branch. Unless the transmembrane resistance is infinite or the intracellular resistance infinitesimal then the current will be progressively divided within each section Δx of the model. As a result the transmembrane potential $\Phi_i - \Phi_e$ will progressively reduce within each section of the model in an exponential fashion. It is possible to show that the transmembrane potential at a distance x along the passively conducting axon is given by [11]:

$$\Phi_i(x) - \Phi_e(x) = V_0 e^{-x/\lambda} \tag{2.3.1}$$

Where λ is the *space constant* (the distance required for the voltage to have dropped by a fraction of $1/e$ from the original value) and is given by:

$$\lambda = \sqrt{r_m/r_i} \tag{2.3.2}$$

Well-designed and properly insulated cables may have a space constant of several hundreds of miles; in contrast the axons of mammals are not so fortunate and have a much lower space constant that may be derived experimentally; for a large myelinated axon from a *xenopus laevis* frog the value of the transmembrane resistance is approximately 600 MΩcm^{-1} and the intracellular resistance is approximately 100 Ωcm^{-1} [12]. Making the assumption that the axon is uninterrupted and does not contain any nodes of Ranvier, using an axonal diameter of 15 μm then the space constant is approximately 4 mm. Table 2.3.1 details the effect of this space constant on an initial transmembrane potential of 1 V; after a distance of 16 mm the transmembrane potential has been attenuated to less than 5% of its original value.

Clearly passive conduction alone is not a viable method by which electrical excitation may be transmitted along the axon for distances greater than a few centimetres. This

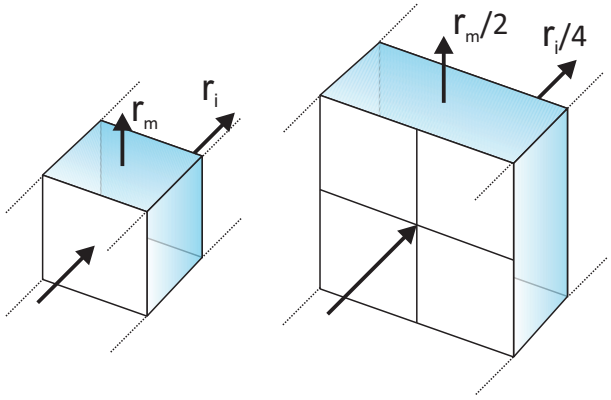


Figure 2.3.3: The effect of cross-sectional area on the intra and transmembrane resistances.

result is a direct consequence of the material properties of the axon that give rise to the comparatively high intracellular resistance as well as the small diameter of axons - typically less than 100 μm [1, 10]. The diameter of the axon has a direct influence on the transmembrane resistance and thus on the space constant. If the diameter of a cylindrical conductor is doubled then the transmembrane resistance r_m is reduced by a factor of two, whilst the longitudinal resistance r_i is reduced by a factor of four. The net effect is a doubling of the ratio $r_m : r_i$ so that the space constant increases by a factor of $\sqrt{2}$. This principle is illustrated in Figure 2.3.3 for a cuboid conductor; the result is analogous to the cylindrical axons.

One solution by which the space constant could be increased is simply to use axons of a larger diameter; however because the net gain in the space constant is only a factor of $\sqrt{2}$ the gains become disproportionate. Consider as an example the axons within the optic nerve, which have a diameter of 10 μm, the space constant is likely to be in the order of a few centimetres, which may very well be sufficient for connecting to the nearby neurons of the lateral geniculate nucleus. The situation is rather different for a sensory neuron within the base of the foot that connects directly to the lumbosacral plexus at the base of the medulla spinalis. If the axon of this neuron had a diameter of 1 cm then its space constant would be approximately 15 cm, so that if you were to shoot yourself in the foot it is likely that this would only be felt at the thigh and below. For a sense of completeness the axon (assuming passive conduction) would need to be about 9 cm in diameter in order to conduct an electrical impulse the required distance, a diameter that is clearly infeasible.

It is clear that passive conduction alone is insufficient for the transmission of electrical impulses over any practical distance. The solution employed by every animal with a nervous system is very much the same solution as devised by Victorian telegraph engineers - the use of *repeaters*. The application of amplifiers or repeaters placed at regular discrete intervals along a cable allows for regeneration of the signal, the effect of which is to negate the attenuation introduction by the previous section of the cable. Conduction with amplification is a process by which energy is injected into the axon; this type of active

conduction is called *propagation*. The axon has the ability to continually regenerate a transmembrane potential so that propagation can proceed over great distance unhindered by the conductive losses described previously.

2.3.2 The Action Potential and Regeneration

If a microelectrode is inserted into the axon and a continuous time recording is made of the transmembrane potential $\Phi_i - \Phi_e$, then it is observed that this potential undergoes a sudden reversal followed by a slower recovery to a resting state. This process can be identically observed at any length along the axon; this impulse is called the *Action Potential* (AP). The first known recording of an AP was made by Hodgkin and Huxley in the late 1930s using an axon from *loligo forbesi* [13]. They postulated that it might be possible to measure the potential between the inside of an axon and the conducting fluid surrounding it. They dissected a large 500 μm neuron from the first stellar nerve and inserted within it a fine cannula. The axon was then submerged in seawater with the lower portion in a layer of insulating mineral oil that allowed for direct electrical stimulation. APs were recorded by connecting an amplifier and recording equipment to one electrode placed in the seawater and the other inside the cannula. In this way Hodgkin and Huxley were for the first time able to make accurate recordings of the transmembrane potential. An example of the AP they recorded is shown in Figure 2.3.4.

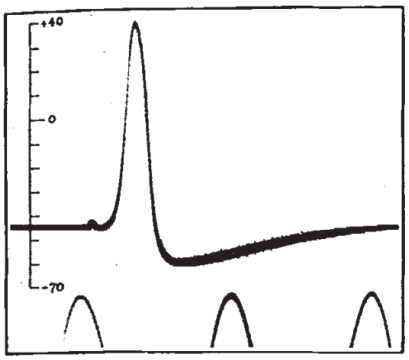


Figure 2.3.4: The AP as recorded from intracellular electrodes placed in an axon from *loligo forbesi*. The time marker is 500 cycles per second and the amplitude scale is in millivolts, reproduced from [13].

If an AP is recorded from a single electrical stimulus at different points along the length of the axon, it can be observed that in fact the amplitude does not decrease as a function of distance (as might be expected for the passive conduction model in Figure 2.3.2) but instead remains constant. Further examination reveals that the morphology of APs is entirely independent of the stimulus which initiated the excitation. So long as the stimulus event is supra-threshold then the same AP is observed travelling along the axon (sub-threshold and no AP is recorded at all). This is an important property of the neuron and is called the *all-or-nothing law* [10]. Clearly there is more at work here than the simple passive conduction predicted with the volume conductor model.

Any connoisseur of explosives will tell you that the rate and temperature at which a black powder fuse burns have no correlation to the match that lit it (causation is not necessarily correlation). The process is entirely chemical and transformative, the thermal energy released by one section of the fuse burning triggers a reaction within the adjoining section,

and so on. The rate and temperature at which the fuse burns are defined purely by the physical and chemical properties of the fuse, so much so that manufacturers specify fuses with highly predictable burn rates - it would be unfortunate if this were not the case. This process represents a suitable analogy to the all-or-nothing law in the axon; the original stimulus induces a longitudinal spread of depolarisation across the cell membrane that conducts passively along the length of the axon. This propagating front, which may be modelled as a current, will give rise to a transmembrane potential. It is this potential that in turn triggers a mechanism within the cell membranes by which a much larger transmembrane potential is produced. This process continues along the length of the axon until the AP has reached the telodendria.

The passive conduction within the axon has already been described and now the regenerative process is considered. It can be observed that the transmembrane resistance drops, albeit briefly, coincident with the regeneration of the AP. The drop in the transmembrane resistance is due to a change in the permeability of the membrane to certain ions, and so it is necessary to understand the nature of this ionic permeability so as to understand better how the AP is regenerated.

2.3.3 Excitable Membranes

The axoplasm or intracellcular fluid of axons has an ionic concentration and properties that are similar to those of most cells, and the axon is surrounded by the extracellular fluid. The chemical differences between these two media are somewhat species specific but broadly speaking the intracellular fluid can be observed to have a good deal more potassium than sodium, conversely the extracellular fluid has more sodium than potassium. The extracellular fluid is remarkably similar in terms of ionic concentrations to seawater; this is not entirely surprising when the evolutionary history of the cell membrane is considered [14].

The difference in the ionic concentrations across the axon membrane is maintained principally via two mechanisms - the overall homeostatic mechanisms that are governed by the renal system determine the characteristics of the extracellular fluid, whilst the intracellular fluid is maintained via particular characteristics of the cell membranes. The most important mechanism behind the balance of this ionic concentration gradient is the sodium-potassium ATP-*ase* or the *sodium pump*. The sodium pump exchanges sodium ions on the outside of the membrane with potassium ions on the inside; this involves the transport of ions against the concentration gradient and so creates a store of potential energy that may be rapidly harvested. In the equilibrium state this ionic concentration gradient produces an transmembrane potential of about -75 mV [13].

As the AP propagates along the axon a localised depolarisation is produced, this depolarisation produces an increase in the permeability of the cell membrane to sodium ions (P_{Na}) that causes the membrane to become further depolarised. As the transmembrane potential

moves towards the equilibrium state where $\Phi_i = \Phi_e$ the permeability of the membrane to potassium ions (P_k) starts to rise and P_{Na} starts to fall causing the transmembrane potential to return to the resting value of -75 mV. There is a delay in the process before P_k begins to rise, during which time the cell membrane is hyperpolarised. The current produced by this depolarisation is spread (via passive conduction) to neighbouring cells and the process is repeated *ad infinitum*. The depolarisation process will only occur when the applied stimulus is supra-threshold for a given axon; the threshold is determined by the ion channel characteristics and varies considerably from one axon to another. The equilibrium state occurs due to a balance between the positive feedback of the sodium channel and the negative feedback of the potassium channel. If the depolarisation (stimulus) is great enough to overcome the stabilising effect of this equilibrium state then the total depolarising process will occur. Below this threshold the balance of concentrations will act to maintain ionic equilibrium and the full scale depolarisation will not occur. This is the origin of the all-or-nothing law and explains why the AP shows identical morphology along the entire length of the axon regardless of the stimulus (assuming that stimulus was supra-threshold).

2.3.4 Conduction Velocity

The speed at which the action potential is propagated along the neuron is constrained by the physical properties of the axon and is far from instantaneous. Intuitively electrical conduction along a simple resistive volume conductor should occur at the speed of light, however when the first transatlantic telegraph cable was laid in 1866 from Newfoundland to Valentia in Ireland, the Victorian engineers were astonished to discover that an electrical impulse took over three seconds to travel the some 2,600 miles. The reason for this reduced conduction velocity is the capacitance between the core and external conductors within the cable and the same situation arises within the axon. Any two electrical conductors that are separated by an insulator will form a capacitor. For a co-axial cable of similar basic construction as the axon the cable capacitance (or transmembrane capacitance) per unit length can be found by:

$$C(x) = \frac{2\pi\epsilon_0\epsilon_r}{ln\left(\frac{D}{d}\right)} \tag{2.3.3}$$

Where d is the outside diameter of the inner conductor, D is the inside diameter of the insulator and $\epsilon_0\epsilon_r$ represents the dielectric constant of the insulator.

The most important observation is that that the capacitance between the two conductors is a function of their relative diameters. For an axon the membrane is typically very thin and so makes a good capacitor; it has a transmembrane capacitance of approximately 1 μFcm^{-2}. The equivalent circuit of Figure 2.3.2 may now be re-drawn to include this capacitance in the form shown in Figure 2.3.5.

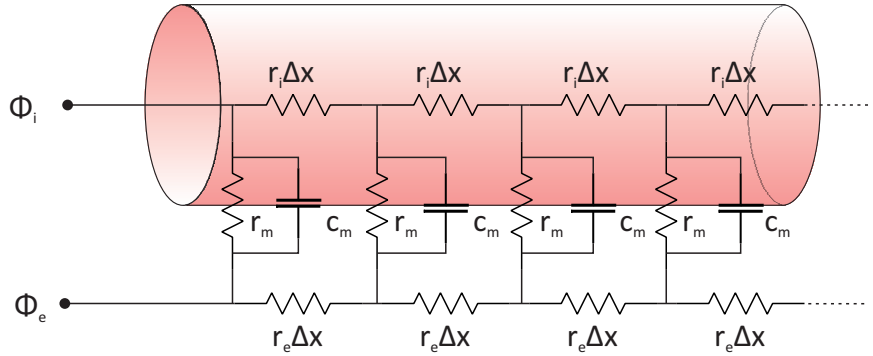


Figure 2.3.5: Revised lumped model parameter including the transmembrane capacitance as well as the longitudinal extracellular resistances.

If a step current is applied to the lumped model shown in Figure 2.3.5 then (neglecting the extracellular resistance r_e) the voltage after one section is given by:

$$V(t) = Ir_m\left(1 - e^{\frac{t}{\tau}}\right) \tag{2.3.4}$$

Where τ is the time constant of the circuit given by $\tau = r_m c_m$ and t represents time.

Table 2.3.2 gives an indication of typical values of r_m and c_m for both the simple cell membrane as well as for the myelin sheath. Typical values of τ are in the order of a few milliseconds and this time constant sets a limit on the speed at which the transmembrane voltage can respond to a localised current flow. Remembering that it is the transmembrane voltage that initiates further depolarisation and propagation, this time constant also limits the overall conduction velocity of the axon. A further process to consider is the time taken for the local sodium permeability to respond to the change in transmembrane potential, which is denoted by T. The conduction velocity for an axon can given by:

$$V = \frac{\lambda}{\tau + T} \tag{2.3.5}$$

Where λ is the time constant and is given by Equation 2.3.2.

Generally speaking $T << \tau$ and so T may be neglected, however there are certain conditions which modulate T substantially enough to affect the conduction velocity along the axon. The primary mechanism is the rate at which the sodium and potassium ions are permitted through the cell membrane, which is temperature specific; higher temperatures will speed up the ionic exchange and thus increase the conduction velocity. The effect is not equal, however, and the potassium channel is more sensitive to temperature variations. As the temperature rises, the potassium channel acts more rapidly and so begins to catch up with the sodium channel, resulting in shorter and smaller APs. This process is self-limiting and in some cold-blooded animals the conduction of APs would cease altogether if the body temperature were to exceed 37 degrees Celsius [3].

Table 2.3.2: Electrical properties of myelin sheath and cell [16].

	Specific Leakage Resistance (Ωcm^2)	**Specific Capacitance (F/cm^2)**
Myelin Sheath	10^5	10^{-8}
Cell Membrane	10^3	10^{-6}

A secondary mechanism that acts on T is the base concentrations of sodium and potassium within the intra and extracellular media, whilst these concentrations are tightly controlled by the homeostatic processes of the renal system they can be modulated by poor dietary intake or severe dehydration. An imbalance of the resting ionic concentrations may lead to an extended recovery phase after depolarisation and ultimately slower conduction of the APs, and in extreme circumstances propagation may cease entirety.

The final mechanism that acts on T is the presence of pharmaceuticals within the blood stream that act directly on the sodium and potassium channels; typically such drugs reduce the speed with which APs are regenerated. One such class of drugs are the tricyclic antidepressants that have a very narrow therapeutic index (the difference between the therapeutic dosage and the toxic dosage) and also block sodium channels leading to a slowing of membrane depolarisation and severely affecting the operation of the myocardium [15].

These mechanisms may be neglected in healthy individuals. The homeostatic regulation of the renal system is highly effective at stabilising the ionic concentrations and temperature changes in the order of a few degrees would be required to noticeably affect conduction velocity, such changes would likely be pathological within themselves and would probably result in death [17].

There are two main factors that influence the conduction velocity by acting on λ & τ:

Diameter - The transmembrane capacitance c_m and resistance r_m are both functions of the axonal diameter. It may be shown from Equation 2.3.3 that the transmembrane capacitance increases with diameter, whislt the transmembrane resistance decreases; these two effects cancel each other out and very little change is observed in τ for different diameters of axon. Recalling from Equation 2.3.2 that λ is proportional to the square root of the ratio of the intracellular and transmembrane resistances and observing that $r_m \propto D^{-1}, r_i \propto D^{-2}$ and $c_m \propto D$ we may deduce that the conduction velocity is proportional to $\sqrt{D}$. Such predictions are confirmed well by experimental data, for example it was reported by Hodgkin and Huxley that for a squid axon held at 18.3 degrees Celsius the conduction velocity Υ roughly followed the relationship:

$$\Upsilon = \sqrt{D} \tag{2.3.6}$$

Myelination - Most axons are myelinated with regular interruptions at every few mil-

limetres at the nodes of Ranvier. Observations of the myelin sheath show that it is made up of a single cell (the *Schwann cell*) that has been wrapped many times around the axon. The effect of the myelin sheath is to thicken the layer of insulation around the axon at all points aside from the nodes of Ranvier. In very much the same way as increasing the axonal diameter, the effect of this extra insulation is to increase the transmembrane resistance and to reduce the transmembrane capacitance (typical values are shown in Table 2.3.2) so the overall effect on the time constant τ is negligible. However because of the relatively high myelin resistance very little current is lost to the extracellular fluid within the internodal region. Propagation in a myelinated axon proceeds from one active node of Ranvier to the next by the virtue of local circuit current. Since there is very little current lost through the transmembrane resistance in the internodal regions the space constant λ is increased and so conduction speeds up. For this reason myelinated axons do not show a square-root relationship for conduction velocity and axonal diameter but rather a linear one. Since the activation jumps from one node to another this type of propagation is said to be *salutatory* (from the Latin *saltare* - to leap or dance).

An important observation of this linear relationship as opposed to the square-root relationship for unmyelinated axons is that the two curves cross at a diameter of about 1 μm. This is the reason why not all of the axons are covered in a myelin sheath. The need for the nodes of Ranvier arises because of the need to replace current that will be lost (albeit slowly) through the still imperfect insulation of the myelin, it is fascinating to observe that the nodes of Ranvier are spaced almost exactly by an order of the space constant [16]. The importance of the myelin sheath is readily made apparent by debilitating conditions such as Multiple Sclerosis, a disease in which the myelin gradually degenerates causing progressive weakness and eventual death.

2.3.5 Neuronal Classification

In the general sense neurons may be classified based on their diameters (and thus conduction velocity); this system was first proposed by Erlanger and Gasser in the 1940s [18]. The precise function of a neuron within the nervous system can usually only be defined by anatomical and physiological examination of its position and connections to surrounding tissue, although Erlanger and Gasser discovered that neurons with different conduction velocities were often associated with particular functions (their observations are detailed in Table 2.3.3). The individual connections within the nervous system of *man* vary considerably from one individual to another, however the basic anatomical structure is consistent. It is possible therefore to determine the function of a particular neuron based both on its conduction velocity and its anatomical position within the structure of the nervous system [19–21]. This is a vitally important result in the design of neuroprostheses - a recording electrode may be placed with accuracy on a specific nerve and if the distribution of conduction velocities within that nerve can be determined then it will be possible not only to identify the function of each axon within the nerve but also to extract directly

Table 2.3.3: The Erlanger-Gasser classification system of nerve fibres [3, 18].

Type	Diameter (μm)	Myelinated	Conduction Velocity (m/s)	Example Function
Aα	8-20	Yes	50-120	Muscle spindle receptors
Aβ	5-12	Yes	30-70	Golgi tendon organ
Aγ	2-8	Yes	10-50	Cutaneous receptors
Aδ	1-5	Moderately	3-30	Touch & pressure receptors
B	1-3	Moderately	3-15	Visceral Sensory
C	<1	No	<2	Nociceptors

the information being transmitted.

2.3.6 The Compound Action Potential

Within a typical fascicle taken from the peripheral nervous system there is a composition of many different types of neuron. There will be a mixture of myelinated and unmyelinated axons that are mixed together, conducting at a wide range of velocities. Table 2.3.3 illustrates the range of possible conduction velocities within the peripheral nervous system of *man* and at the extreme the axon may propagate AP as fast as 120 m/s or as slow as less than 1 m/s. A consequence of this mixture of velocities is that if a fascicle, or a nerve, is stimulated electronically at some point and a recording is made at some other point along the same nerve then the observed response is a *compound action potential* (CAP). The CAP is the result of the superposition of many different APs all occurring at roughly the same time, the further away the recording is made from the stimulation electrode the more the individual APs are spread out. This result is a direct consequence of the range of different conduction velocities. It is often observed that each fascicle actually supports a number of *populations* of conduction velocity, if the distribution of axon diameters was plotted as a histogram there would be distinct peaks that correspond to each population. The peaks observed in the CAP therefore give a rudimentary spectrum of the conduction velocities within the fascicle.

2.4 Methods of Signalling

It is important to understand how neurons transmit information around the body. We live in an analogue world and the nervous system must be capable of communicating a large range of different sensory and motor commands that represent these analogue values. Assuming passive conduction with relatively short nerves then it might have been possible to encode an analogue stimulus in terms of the AP amplitude or width. However, the all-or-nothing law associated with active propagation prohibits such a mechanism; the law produces APs that are the same duration and amplitude regardless of the original stimulus. Consider then a more digital form of communication whereby information is encoded by discrete pulses (or APs). The all-or-nothing law will faithfully reproduce the APs as they travel along the axon in an effectively lossless communication channel. It has

been found that the nervous system transmits information in very much the same way as traditional broadcast radio, in a method akin to *frequency modulation*. The number of APs transmitted along a nerve per second is directly correlated to the amplitude of the stimulus [22].

There are many advantages to using frequency modulation to transmit information, firstly there is no requirement for a clock signal; the system is robust in the presence of additive noise and it is possible to represent a wide range of analogue values. In theory a single axon has a finite range starting from zero activity (DC) up to the limit where APs are occurring *back-to-back*. A typical AP is approximately 1 ms in duration [23], and so the maximum number of APs per second is around one thousand. Given this fundamental digital coding scheme the dynamic range of representable values is approximately 60 dB. This is quite a costly method of communication as the constant depolarisation of the cell membrane requires the transport of both sodium and potassium ions using ATP, but it is very robust. A dynamic range of 60 dB may not be sufficient for encoding the full range of possible values and so the nervous system has the ability to increase the available dynamic range; this is in part the reason why there are often so many nerves attached to a single muscle or sensor. Each neuron has a different threshold from its neighbours; this provides a method for fine grading the level of stimulus and is known as recruitment. As each neuron has a different threshold it is found that as the level of stimulus is increased, it is not only that a single fibre produces more APs but that more neurons start to fire. This concept applies equally well to both motor and sensory fibres; a good example is the nerve from the vestibular apparatus in which the neurons are all found to have different thresholds, so increasing stimulation leads to more of them firing simultaneously [3].

The process by which a physical stimulus is converted into a burst of APs is not straightforward, suffice to say that if the receptor channels of the soma are held open then the membrane potential will move towards zero and at some point must cross the threshold which triggers an AP within the axon. The usual sequence in changes of permeability will then occur and the membrane potential will return to the resting state, if the receptor channels are still held open then the potential will once again move towards zero creating another AP as the membrane potential crosses the threshold. Impulses will continue to be generated so long as the receptor channels are held open by some stimulus, the greater the short circuiting current (or stimulus size) the faster the cell will depolarise and the faster an AP will be generated. Thus the frequency of firing will be dependent on the size of short circuiting current, or the number of receptor channels that are open. Under ideal conditions the firing frequency is a simple, often linear, function of the applied current [24].

2.4.1 Refractory Period

If a neuron is repeatedly stimulated with a train of impulses while reducing the inter-pulse interval then there comes a point at which the apparent threshold for successful stimulation increases. If the inter-pulse interval continues to decrease, a limit is reached at which point the axon cannot be stimulated regardless of the magnitude of the stimulation current applied. In other words, the neuron has a refractory period immediately following the generation of AP during which time it is impossible to elicit further APs. The period during which the neuron may be stimulated - but only by using larger stimulation currents than usual - is called the relative refractory period. These refractory periods arise from the properties of the sodium channels, just after an AP has occurred the sodium mechanism is unresponsive, making the membrane absolutely stable to any stimulus. This refractory period also serves a vital functional role: it prevents the propagation of APs in both directions simultaneously. Referring to the assumptions made in the simple volume conductor model of Figure 2.3.2, the longitudinal currents would flow equally in both directions along the axon. It is therefore essential that the region over which an AP has just passed should not be permitted to regenerate the same AP immediately afterwards. The refractory period prevents just such an occurrence.

2.4.2 Accommodation

The rates at which sodium and potassium ions can move through the channels within the membrane are not identical and in fact sodium is able to move through the channel significantly more rapidly than potassium. One of the consequences of this is that the neuron is more sensitive to rapid changes in the transmembrane potential as the sodium channel can exchange ions before the much slower potassium channel has a chance to move towards an equilibrium state. This effect is called *accommodation*; the neuron is more readily stimulated by fast depolarisations than it is by slow depolarisations.

Considering stimulation of the neuron with a staircase-like pattern of increasing current; P_k increases cumulatively with each step, whereas P_{Na} does not because it is only transitory. The transient increases in P_{Na} will steadily decline with increasing stimulation current because of the increasing degree of sodium channel inactivation. The more gradually a neuron is depolarised the more the sodium/potassium balance is pushed towards potassium, raising the effective stimulation threshold required for AP generation. If the neuron is depolarised slowly enough there will come a point at which P_{Na} is never large enough relative to P_k for the neuron to generate an AP at all. The membrane is said to have *fully accommodated*. The mechanism of accommodation can be an important determinant of the way in which sensory receptors respond to slowly changing stimuli.

2.5 Recording Techniques

2.5.1 Electrodes for Stimulation and Recording

In recent years there have been significant advances in the design of neuroprostheses intended to restore functions that have previously been lost or otherwise impaired through spinal cord injury, specifically the restoration of motor functions such as hand grasping and the partial restoration of bladder control [25–27]. *Functional electrical stimulation* (FES) of the nervous system requires some form of control and in current systems it is often the user who must activate the stimulator. For example a commercial bladder stimulator requires the user to measure fluid intake and outtake, manually activating the bladder stimulator based on the theoretical fluid retained [28, 29]. It would be a significant benefit to a patients quality of life if a recording system could interface to the nerves that innervate the bladder and provide direct feedback about the true level of retained fluid. Preliminary work has shown that information extracted from the *electroneurogram* (ENG) of a sensory nerve can, once suitably filtered and otherwise processed, be used to provide a significant improvement in the performance of a functional neural prosthesis [26].

When considering the method by which to extract information from the nervous system there are three key requirements and limitations that must be clearly defined and may be used as a metric for evaluation,

1. **Surgical Invasiveness** - There are a number of practical advantages to minimising the invasiveness of the surgery required to implant a viable recording device into the human body. Firstly, the avoidance of large wounds should lead to a reduction in both healing time and postoperative pain; this should in turn lead to earlier hospital discharge [30]. Secondly, there is a risk involved with any level of surgical procedure, but this risk can be minimised if the implanting of a device does not require any destruction or alteration of the nervous system.

2. **Information Extraction** - In order to be a viable candidate for implantation any system must be capable of extracting information from the nervous system. This information must be extracted automatically so as to avoid the need for regular visits to hospitals or clinics. It may not be necessary to continually extract information however and biological or electrical triggers may be used to start processing.

3. **Stability in Chronic Implantation** - Closely related to the surgical invasiveness of an electrode is the stability of the nerve and recording arrangement over long periods of time [31, 32]. The fragility of the nerve limits the type of electrodes that can be used for chronic implantation as the nerve is highly sensitive to constriction or surface damage.

2.5.2 Percutaneous Electrodes

In situations where a target nerve lies relatively close to the surface of the skin, a simple wire like electrode known as a *percutaneous*, or through skin, electrode can be used. Percutaneous electrodes are inserted through the skin and into a nerve guided predominantly by anatomical landmarks but also by medical imaging such as ultrasound [33]. When locating a target nerve it is possible to use a stimulating electrode to identify a target area based on a trial and error basis; a stimulation pulse is applied to different anatomical areas around the nerve until the desired motor or sensory response is achieved. Percutaneous electrodes are placed proximally to the nerve but not in direct contact so stimulation will result in the recruitment of surrounding muscle and soft tissue as well as the nerve fibre itself. For this reason they are only used for stimulation, as recorded signals from muscles (*electromyogram* (EMG)) would be of much greater amplitude than any signals from the nerve fibre, and would therefore mask the nerve signals.

The primary advantage of percutaneous electrodes is minimal surgical invasiveness, which allows for relatively safe experimentation for minor disorders where there is no requirement for more radical and invasive procedures. Percutaneous electrodes have been used in the restoration of functional hand grip in paralyzed individuals, with up to 88% of implanted electrodes surviving for over six months allowing for outpatient use of an upper extremity neuroprosthetic system [34]. Percutaneous electrodes have also been used for pain relief in patients by direct stimulation of the peripheral nerves [35].

2.5.3 Epineurial Electrodes

Epineurial electrodes are annular discs that are surgically attached to the outside of the target nerve by sewing them to the epineurium. This provides a direct connection to the nerve with a relatively high selectivity for stimulation purposes, however the electrode ring has no shielding from the surrounding tissue and so without modification is unsuitable for neural recording purposes as the low amplitude signals from the nerve fibre are masked by the much larger muscle signals. Epineurial recording of APs has been reported for determining conduction velocity; however a silicone sheet was required to isolate the electrodes from the surrounding tissue and thus improve the signal quality [36]. The epineurial electrode only rests on the surface of the nerve; it does not penetrate or in any way constrain the nerve fibre. This makes it particularly well suited to chronic implantation and epineurial electrodes have been chosen over cuff electrodes because of the risk of damage from the more constraining cuff electrode [37]. When using multiple epineurial electrodes placed around the circumference of a large nerve fibre it is possible to recruit targeted axons within the nerve, this allows for selective stimulation that can reduce the effects of fatigue caused by prolonged stimulation by alternating which fibres are stimulated. The safety and long term stability of epineurial electrodes has been well documented [37]. A more basic epineurial electrode can be formed by simply resting the nerve on an electrode

that has been shaped into a hook, this method however is only suitable for acute experimentation as it requires partial removal of the nerve from the body. Hooks are used later in this study for the acquisition of ENG from *rat* in Chapter 6.

2.5.4 Circumferential Cuff Electrodes

Circumferential cuff electrodes consist of an insulating tube made of either silicone or polyimide, inside which is fitted a number of annular electrodes typically made from platinum [38], although other types are avaliable. Surgical implantation is achieved by opening the cuff along its length and wrapping it around the nerve, the cuff is then closed trapping the nerve inside. Cuff electrodes have a very successful history as chronically implanted recording [39] and stimulation electrodes [40] and they have been used for both commercial and research devices for over four decades [41]. The purpose of the cuff design is to maintain both a good physical contact with the nerve and to provide a means of isolation and shielding from interference due to nearby muscle activity. Although the cuff does provide some measure of shielding from noise sources such as EMG there is still scope for these currents to flow into the ends of the cuff structure and contaminate any recordings. It has been demonstrated that the use of a *tripolar* recording configuration (wherein the end electrodes are grounded) significantly reduces the effect of external current sources such as EMG [42, 43]. Topologically selective recording from nerves using a cuff electrode is difficult. The recorded potential at any given time is a weighted sum of all active sources within the fibre, while fascicles that are closer to the surface of the fibre (and thus proximal to the recording electrodes) will produce a larger response, it is possible that a much larger fascicle in the middle of the fibre will hamper selective recording.

It has been shown that if the spacing between the electrodes is uniform then it is possible to selectively record neural signals based on the speed with which they propagate along the fibre [44–46]. If multiple electrodes are used, and the sampling time is known, the recordings can be averaged after adding a delay to each successive electrode, depending on the expected velocity of the fibre group of interest. This approach will be considered in more detail in Section 2.6.

2.5.5 Interfascicular Electrodes

Typically electrodes are placed either proximally or in direct contact with the epineurium, thus minimising any possible destruction to the fascicles and maximising the possibility of chronic implantation. Electrodes that are placed inside a peripheral nerve but outside of the fascicle have been developed in an effort to increase the selectivity available with respect to traditional extraneural electrodes and also to increase the *signal-to-noise-ratio* (SNR). The goal of these *interfascicular* electrodes is to stimulate or record only from the fascicles that are in direct contact with the electrode, however there is still a high probability that stimulation will recruit nearby fascicles [47]. Several interfascicular electrodes

may be implanted to increase the level of topological selectivity; typically the electrodes are bluntly inserted into the fibre penetrating the epineurium without compromising the integrity of the perineurium [47].

An example of an interfascicular electrode is the *Slowly penetrating interfascicular electrode* (SPINE), the SPINE consists of a silicone tube with blunt electrodes extending radially into the lumen of the tube that penetrates the epineurium. The electrodes are pushed into the epineurium after implantation by the elasticity of the interface, essentially separating the nerve into four electrically isolated chambers. While there have been no reports of chronic implantation, the histological cross-sections showed that the SPINE rearranges the epineurium and penetrates deep within a multifascicular nerve but did not disrupt the perineurium [48].

A less invasive interfascicular electrode is the *flat interface nerve electrode* (FINE) [49]. The FINE is a book like structure formed from silicone elastomer that is placed around a fascicle and sutured closed, applying a small amount of pressure that over a period of a few days will spread the fascicle over a larger surface area, thus reducing the electrode to axon separation distance. The electrodes within the FINE are made from platinum foil and studies have shown that if the pressure applied is low enough then little to no damage is incurred by the fascicle, in contrast to the SPINE that penetrates the epineurium.

2.5.6 Intrafascicular Electrodes

Intrafascicular electrodes are wire like structures that penetrate the perineurium and are inserted inside the fascicles. They aim to provide fascicular or subfascicular stimulation and recording [50–52]. *Longitudinally implanted intrafascicular electrodes* (LIFEs) provide a means of interfacing to a subset of axons within a multi-fascicle peripheral nerve. LIFEs are constructed from a Teflon coated core of platinum-iridium (Pt-Ir) or stainless steel and have a diameter of 25 μm $-$ 50 μm. The active recording or stimulation zone is a short section (250 μm $-$ 1,500 μm) which has been stripped of insulation [41]. The stiffness of metal wires can lead to motion of the electrode which elicits fibrous encapsulation and degrades performance over time, for this reason flexible polymer filaments are now preferred [41].

A similar type of intrafascicular electrode is the *transverse intrafascicular multichannel electrode* (TIME) [50]. The TIME is similar in basic construction to the LIFE but is implanted transversely as opposed to longitudinally in order to selectively recruit a greater spread of axons within different fascicles. The TIME electrode has been demonstrated in acute experiments but has yet to be examined in a chronic study.

LIFE and TIME provide good selectivity for both stimulation and recording, due to the proximity of the electrodes to individual nerves it is possible to record from a single axon. The most important factor in determining which nerve is recorded is the nerve

to electrode distance; fibres which are proximal to the electrode are most likely to be recorded [53]. Intrafascicular electrodes are by nature highly invasive interfaces; they penetrate the perineurium which is the natural protecting sheet of the nerve. The sharp incision procedure and subsequent movement of the electrode within the fascicle may create lesions. The flexibility of polymer based electrodes goes some way towards alleviating the problems associated with movement; however the build-up of connective tissue as an immune response can degrade the performance of the device in chronic implantation [54].

2.5.7 Penetrating Array Electrodes

Penetrating array electrodes are also intrafascicular in nature, but they follow a radically different design. Instead of using a fine wire like structure like a LIFE, penetrating arrays consist of an array of rigid, needle like structures as demonstrated in the Utah Array.

Originally designed for cortical implantation, the array is bluntly inserted into the nerve at high velocity using a pneumatic device, piercing both the epineurium and the perineurium [55]. Like the LIFE structure, an electrode array has the ability to record at the fibre level and the ability to stimulate at the fascicle or sub-fascicle level (assuming an electrode pierces a fascicle). Typically one hundred needle like structures form the array. Unlike a LIFE structure the needles cannot be fabricated from a flexible polymer, so the device remains rigid and must be securely attached to the nerve using a cuff or suture [56]. Damage to the nerve can be caused by the traumatic method of insertion, constriction by the cuff or fixation device and by nerve movement around the electrodes. Signs of severe nerve trauma following implantation as well as long term axonal damage have been reported in *cat* [56].

2.5.8 Regenerative Electrodes

Regenerative electrodes are designed to interface with each axon individually, allowing for the stimulation and recording of single APs within individual axons. The nerve is first transected; regrowth is then supported via a mechanical structure within the electrode. The interface typically consists of a tube structure with a perforated disk in the middle, which contains electrodes in each hole. The ends of the transected nerve are placed into each end of the tube and allowed to regrow through the perforated disk, creating a highly selective interface. In the most ideal case regenerative electrodes would aim to interface to each individual axon. This is not feasible however; it is currently impractical to produce sufficiently small electrodes and growth of multiple axons through the same hole cannot be prevented [41].

It has been demonstrated that neurons will regenerate through a sieve-like structure, and the sieves are functional as both recording and stimulation devices, although difficulties

have been encountered when used for chronic implantation, likely due to incomplete or constrained regeneration of the axons [57, 58]. Total regeneration of the nerve fibre however has not yet been proven and long term implantation suggests that this interface inflicts repeated long term damage to the nerve [57]. A likely candidate for the cause of this damage is the constrictive forces of the electrodes; when the axons begin to regenerate they will be very thin and it is probable that many axons will grow through the same hole, but as regeneration is completed and the myelin sheath is re-established they increase in diameter leading to constriction and eventual damage [57].

2.5.9 Microchannel Electrodes

Microchannel electrode arrays are a recent advancement and aim to improve the SNR obtained for acute recordings by subdividing axons at the fascicular level. The interface is formed from *polydimethylsiloxane* (PDMS) which provides a level of flexibility and allows the interface to twist and conform to the nervous tissue during implantation. The reference devices consist of four 5 mm long channels with a cross sectional area of 100 μm by 100 μm. A main recording electrode and two reference electrodes are inserted into each channel to facilitate a tripolar recording. This type of electrode structure requires that the axons be cut and teased through the channels using sutures and has been successfully demonstrated recording from $S1$ and $L6$ dorsal roots of anaesthetized *rats* [59, 60].

2.5.10 Comparisons

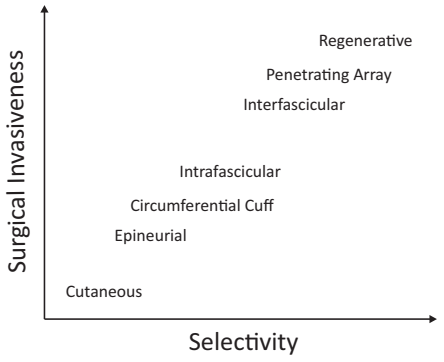


Figure 2.5.1: Illustration of the relationship between selectivity and surgical invasiveness.

As the interface between the peripheral nervous system and an artificial control system, the nerve electrode is a key component of any neural prosthetic device. The performance of the interface will likely define the functionality of the overall system, the level of selectivity must be made as high as possible in order to extract the maximum amount of information from the nervous system. An interface to each individual axon would provide the highest possible level of selectivity, but this is rarely a practical approach. There is a trade-off between selectivity and surgical invasiveness when choosing which electrode configuration to use, high selectivity tends to imply a highly invasive interface, as the electrode to axon distance must be reduced, which may not be suitable for chronic implantation. Figure 2.5.1 represents the relationship between selectivity and invasiveness for the nerve interfaces discussed previously. When choosing an interface it is important to consider the level of selectivity that is required for a functional prosthesis, as well as the likely benefit that will become the patient and if these

benefits justify the risks involved.

Cuff electrodes provide a stable and safe interface that has been well documented in clinical trials on humans, they provide a functional level of selectivity but importantly they have the potential for artefact rejection when used in a tripolar configuration that is not available to any other interface that does not penetrate the nerve. They can be chronically implanted and do not interrupt the functional operation of the nerve, making them a promising candidate for long term neural interfaces. If the native selectivity of a cuff could be increased, either via electrode design and layout or through signal processing, then even more information could be extracted for the same level of surgical invasiveness. The use of velocity selective recording with electrode cuffs has already demonstrated a mechanism for artificially increasing the selectivity and it is this approach that will be investigated further.

2.6 Multiple Electrode Cuffs

Multiple electrode cuffs (MECs) have been used extensively for both chronic stimulation [61, 62] and recording [40]. There are several basic approaches to the manufacture of MECs. These include wrapping wires around the nerve and bonding them in place with a dental impression compound, threading an open spiral assembly with wire contacts along the nerve and the more recent use of photolithographic metallization and machining to produce a thin-film pattern on a flexible polymer substrate [40]. Any of these methods allows for relative ease of implantation, either the cuff is slit and wrapped around the nerve or the cuff is self-closing, this allows surgical placement without damage to the nerve [63].

Self-sizing cuffs are available which are designed to expand and contract in response to any inflammatory action of the nerve, they have been used successfully to stimulate the optic nerve of patients with *retinitis pigmentosa*; a condition where the photoreceptor cells of the eyes are destroyed while the ganglion cells and their axonal processes remain [64].

The length of the cuff may place limitations on possible sites for implantation, however to date nerve cuffs have been successfully used at sites in the limbs [27] and on the nerve that innervate the urinary bladder [65]. Typically only one signal is available from each cuff if a single electrode is used and so the information than can obtained is limited. The signal recorded at the cuff will be the result of the complex interactions and superposition of all the active fibres within the nerve, typically the nerve will contain a great number of both afferent and efferent fibres so the recording of a single signal represents a great loss of information.

Some functionality can be extracted from single recordings if the nerve is transected distal to the point of sensory interest, thus removing a large amount of unwanted neural traffic from the nerve, but this method is undesirably destructive [66].

The addition of multiple electrodes to a cuff is not a recent idea. Extra electrodes have been used for selective stimulation in *cat* by adding one or more anodes to steer the stimulation current away from non-target areas of the nerve [64]. Using such methods it is possible to recruit selectively small areas of the nerve. Adding more electrodes along the length of the cuff opens up the possibility of recording a wide range of different conduction velocities, the greater the electrode spacing the higher the maximum recordable velocity for a given sample rate. This technique, the discrimination of neural traffic based on conduction velocity, is called *velocity selective recording* (VSR).

2.7 Conclusions and Requirements for a Practical Recording System

This thesis concentrates on the application of real time methods such as VSR to the analysis of the ENG. Real time recording is preferable, if not essential, in order to provide the level of feedback required for a FES based neuroprostheses. This limitation has so far prevented the use of the traditional offline based processing methods such as spike sorting. Not only must the recording system function in real time without training but it must also be chronically implantable, a requirement that is readily fulfilled using an electrode structure such as the MEC. For these reasons the work in this thesis will concentrate on the extension of VSR to naturally occurring ENG recorded with a chronically implantable electrode structure.

References

[1] P. Williams, R. Warick, M. Dyson, and L. Bannister, "Major Divisions of the Nervous System," in *Gray's Anatomy*, ch. 7, p. 919, 37 ed., 1989.

[2] J. R. Heiman, "A psychophysiological exploration of sexual arousal patterns in females and males.," *Psychophysiology*, vol. 14, pp. 266–74, May 1977.

[3] R. Carpenter, *Neurophysiology*. Arnold Publishers, fourth ed., 2003.

[4] F. A. C. Azevedo, L. R. B. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. L. Ferretti, R. E. P. Leite, W. Jacob Filho, R. Lent, and S. Herculano-Houzel, "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.," *The Journal of comparative neurology*, vol. 513, pp. 532–41, Apr. 2009.

[5] M. P. Coleman and M. R. Freeman, "Wallerian degeneration, wld(s), and nmnat," *Annual review of neuroscience*, vol. 33, pp. 245–67, Jan. 2010.

[6] A. Taylor, "The mechanism of injury to the spinal cord in the neck without damage to the vertebral column," *Journal of Bone & Joint Surgery, British Volume*, no. 33, 1951.

[7] R. Bakshi, Z. A. Shaikh, R. S. Miletich, D. Czarnecki, J. Dmochowski, K. Henschel, V. Janardhan, N. Dubey, and P. R. Kinkel, "Fatigue in multiple sclerosis and its relationship to depression and neurologic disability," *Multiple Sclerosis*, vol. 6, pp. 181–185, June 2000.

[8] B. Hille, *Ion channels of excitable membranes*. 2001.

[9] T. Matsuyama, M. Mackay, and R. Midha, "Peripheral nerve repair and grafting techniques: a review," *Neurologia medico-chirurgica*, vol. 40, pp. 187–199, 2000.

[10] J. Nicholls, A. Martin, and B. Wallace, *From Neuron To Brain*. Sinauer Associates, third ed., 1998.

[11] A. L. Hodgkin and W. A. H. Rushton, "The Electrical Constants of a Crustacean Nerve Fibre," *Proceedings of the Royal Society B: Biological Sciences*, vol. 133, pp. 444–479, Dec. 1946.

[12] K. W. Altman and R. Plonsey, "Analysis of the longitudinal and radial resistivity measurements of the nerve trunk.," *Annals of biomedical engineering*, vol. 17, pp. 313–24, Jan. 1989.

[13] A. Hodgkin and A. Huxley, "Action potentials recorded from inside a nerve fibre," *Nature*, vol. 144, pp. 710–711, 1939.

[14] W. A. H. Rushton, "A theory of the effects of fibre size in medullated nerve," *The Journal of physiology*, vol. 115, no. 1, pp. 101–122, 1951.

[15] G. W. Kerr, "Tricyclic antidepressant overdose: a review," *Emergency Medicine Journal*, vol. 18, pp. 236–241, July 2001.

[16] R. Plonsey and R. Barr, "Impulse Propogation," in *Bioelectricity: A Quantitative Approach*, ch. 6, pp. 155–185, Springer, third ed., 2007.

[17] G. Schalow, G. Zäch, and R. Warzok, "Classification of human peripheral nerve fibre groups by conduction velocity and nerve fibre diameter is preserved following spinal cord lesion," *Journal of the autonomic nervous system*, vol. 1838, no. 6, 1995.

[18] H. Gasser, "The classification of nerve fibers.," *Ohio Journal of Science*, vol. 41, no. 3, pp. 145–159, 1941.

[19] G. M. Manzano, L. M. P. Giuliano, and J. M. Nóbrega, "A brief historical note on the classification of nerve fibers.," *Arquivos de neuro-psiquiatria*, vol. 66, pp. 117–9, Mar. 2008.

[20] B. Andrew and N. Part, "Properties of fast and slow motor units in hind limb and tail muscles of the rat," *Experimental Physiology*, pp. 213–225, 1972.

[21] P. Williams, R. Warick, M. Dyson, and L. Bannister, "Structure of Neurons," in *Gray's Anatomy*, ch. 7, pp. 875–876, 37 ed., 1989.

[22] G. Stuart, J. Schiller, and B. Sakmann, "Action potential initiation and propagation in rat neocortical pyramidal neurons," *The Journal of Physiology*, vol. 505, pp. 617–32, Dec. 1997.

[23] B. Frankenhaeuser and A. Huxley, "The action potential in the myelinated nerve fibre of Xenopus laevis as computed on the basis of voltage clamp data," *The Journal of Physiology*, pp. 302–315, 1964.

[24] R. Granit, D. Kernell, and G. Shortess, "Quantitative aspects of repetitive firing of mammalian motoneurones, caused by injected currents," *The Journal of physiology*, pp. 911–931, 1963.

[25] M. Haugland, A. Lickel, J. Haase, and T. Sinkjaer, "Control of FES thumb force using slip information obtained from the cutaneous electroneurogram in quadriplegic man.," *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 7, pp. 215–27, June 1999.

[26] M. Haugland, J. Hoffer, and T. Sinkjaer, "Skin contact force information in sensory nerve signals recorded by implanted cuff electrodes," *IEEE Transactions on Rehabilitation Engineering*, vol. 2, pp. 18–28, Mar. 1994.

[27] M. Haugland and T. Sinkjaer, "Cutaneous whole nerve recordings used for correction of footdrop in hemiplegic man," *IEEE Transactions on Rehabilitation Engineering*, vol. 3, no. 4, pp. 307–317, 1995.

[28] G. S. Brindley, C. E. Polkey, and D. N. Rushton, "Sacral anterior root stimulators for bladder control in paraplegia.," *Paraplegia*, vol. 20, pp. 365–81, Dec. 1982.

[29] P. Dasgupta, C. Haslam, R. Goodwin, and C. J. Fowler, "The 'Queen Square bladder stimulator': a device for assisting emptying of the neurogenic bladder.," *British journal of urology*, vol. 80, pp. 234–7, Aug. 1997.

[30] K. Fuchs and B. Jaffray, "Minimally invasive surgery.," *Archives of disease in childhood*, vol. 90, pp. 537–42, May 2005.

[31] W. M. Grill, S. E. Norman, and R. V. Bellamkonda, "Implanted neural interfaces: biochallenges and engineered solutions.," *Annual review of biomedical engineering*, vol. 11, pp. 1–24, Jan. 2009.

[32] C. A. Miller, P. J. Abbas, J. T. Rubinstein, B. K. Robinson, A. J. Matsuoka, and G. Woodworth, "Electrically evoked compound action potentials of guinea pig and cat: responses to monopolar, monophasic stimulation.," *Hearing research*, vol. 119, pp. 142–54, May 1998.

[33] M. A. Huntoon and A. H. Burgher, "Review of ultrasound-guided peripheral nerve stimulation," *Techniques in Regional Anesthesia and Pain Management*, vol. 13, pp. 121–127, July 2009.

[34] W. Memberg, P. Peckham, G. Thrope, M. Keith, and T. Kicher, "An analysis of the reliability of percutaneous intramuscular electrodes in upper extremity FNS applications," *IEEE Transactions on Rehabilitation Engineering*, vol. 1, pp. 126–132, June 1993.

[35] R. J. Mobbs and A. Lazarro, "Stimulation of the medial plantar nerve for complex regional pain syndrome.," *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia*, vol. 17, pp. 1421–2, Nov. 2010.

[36] B. Murphy, C. Krieger, and J. Hoffer, "Chronically implanted epineural electrodes for repeated assessment of nerve conduction velocity and compound action potential amplitude in rodents," *Journal of Neuroscience Methods*, vol. 132, pp. 25–33, Jan. 2004.

[37] G. Creasey, J. Elefteriades, A. DiMarco, P. Talonen, M. Bijak, W. Girsch, and C. Kantor, "Electrical stimulation to restore respiration.," *Journal of rehabilitation research and development*, vol. 33, pp. 123–32, Apr. 1996.

[38] J. S. Knutson, G. G. Naples, P. H. Peckham, and M. W. Keith, "Electrode fracture rates and occurrences of infection and granuloma associated with percutaneous intramuscular electrodes in upper-limb functional electrical stimulation applications.," *Journal of rehabilitation research and development*, vol. 39, no. 6, pp. 671–83, 2002.

[39] L. Andreasen and J. Struijk, "Signal strength versus cuff length in nerve cuff electrode recordings," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 9, pp. 1045–1050, 2002.

[40] G. E. Loeb and R. A. Peck, "Cuff electrodes for chronic stimulation and recording of peripheral nerve activity.," *Journal of neuroscience methods*, vol. 64, pp. 95–103, Jan. 1996.

[41] X. Navarro, T. B. Krueger, N. Lago, S. Micera, T. Stieglitz, and P. Dario, "A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems.," *Journal of the peripheral nervous system : JPNS*, vol. 10, pp. 229–58, Sept. 2005.

[42] L. N. S. Andreasen and J. J. Struijk, "Model-based evaluation of the short-circuited tripolar cuff configuration.," *Medical & biological engineering & computing*, vol. 44, pp. 404–13, May 2006.

[43] L. Andreasen and J. Struijk, "Artefact reduction with alternative cuff configurations," *IEEE TRansactions on Biomedical Engineering*, vol. 50, no. 10, pp. 1160–1166, 2003.

[44] K. Yoshida, G. A. M. Kurstjens, and K. Hennings, "Experimental validation of the nerve conduction velocity selective recording technique using a multi-contact cuff electrode," *Medical engineering & physics*, vol. 31, pp. 1261–1270, Dec. 2009.

[45] J. Taylor, N. Donaldson, and J. Winter, "Multiple-electrode nerve cuffs for low-velocity and velocity-selective neural recording," *Medical & Biological Engineering*, vol. 42, pp. 634–643, Sept. 2004.

[46] M. Schuettler, V. Seetohul, J. Taylor, and N. Donaldson, "Velocity-selective recording from frog nerve using a multi-contact cuff electrode.," *Conference Proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 2962–5, Jan. 2006.

[47] P. Veltink, J. V. Alste, and H. Boom, "Multielectrode intrafascicular and extraneural stimulation," *Medical and Biological Engineering and Computing*, no. January, pp. 19–24, 1989.

[48] D. J. Tyler and D. M. Durand, "A slowly penetrating interfascicular nerve electrode for selective activation of peripheral nerves.," *IEEE transactions on rehabilitation engineering*, vol. 5, pp. 51–61, Mar. 1997.

[49] D. Tyler and D. Durand, "Functionally selective peripheral nerve stimulation with a flat interface nerve electrode," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 10, pp. 294–303, Dec 2002.

[50] T. Boretius, J. Badia, A. Pascual-Font, M. Schuettler, X. Navarro, K. Yoshida, and T. Stieglitz, "A transverse intrafascicular multichannel electrode (TIME) to interface with the peripheral nerve.," *Biosensors & bioelectronics*, vol. 26, pp. 62–9, Sept. 2010.

[51] J. Badia, T. Boretius, D. Andreu, C. Azevedo-Coste, T. Stieglitz, and X. Navarro, "Comparative analysis of transverse intrafascicular multichannel, longitudinal intrafascicular and multipolar cuff electrodes for the selective stimulation of nerve fascicles.," *Journal of neural engineering*, vol. 8, June 2011.

[52] P. H. Veltink, B. K. van Veen, J. J. Struijk, J. A. N. Holsheimer, and H. B. K. Boom, "A modeling study of nerve fascicle stimulation.," *IEEE transactions on bio-medical engineering*, vol. 36, pp. 683–92, July 1989.

[53] K. Yoshida and R. B. Stein, "Characterization of signals and noise rejection with bipolar longitudinal intrafascicular electrodes.," *IEEE transactions on bio-medical engineering*, vol. 46, pp. 226–34, Feb. 1999.

[54] T. Lefurge, E. Goodall, K. Horch, L. Stensaas, and A. Schoenberg, "Chronically implanted intrafascicular recording electrodes.," *Annals of biomedical engineering*, vol. 19, pp. 197–207, Jan. 1991.

[55] K. E. Jones, P. K. Campbell, and R. A. Normann, "A glass/silicon composite intracortical electrode array.," *Annals of biomedical engineering*, vol. 20, pp. 423–37, Jan. 1992.

[56] A. Branner, R. B. Stein, E. Fernandez, Y. Aoyagi, and R. A. Normann, "Long-term stimulation and recording with a penetrating microelectrode array in cat sciatic nerve.," *IEEE transactions on bio-medical engineering*, vol. 51, pp. 146–57, Jan. 2004.

[57] N. Lago, D. Ceballos, F. J. Rodríguez, T. Stieglitz, and X. Navarro, "Long term assessment of axonal regeneration through polyimide regenerative electrodes to interface the peripheral nerve.," *Biomaterials*, vol. 26, pp. 2021–31, May 2005.

[58] R. M. Bradley, X. Cao, T. Akin, and K. Najafi, "Long term chronic recordings from peripheral sensory fibers using a sieve electrode array.," *Journal of neuroscience methods*, vol. 73, pp. 177–86, May 1997.

[59] E. Delivopoulos, D. J. Chew, I. R. Minev, J. W. Fawcett, and S. P. Lacour, "Concurrent recordings of bladder afferents from multiple nerves using a microfabricated PDMS microchannel electrode array.," *Lab on a chip*, vol. 12, pp. 2540–51, July 2012.

[60] D. J. Chew, L. Zhu, E. Delivopoulos, I. R. Minev, K. M. Musick, C. A. Mosse, M. Craggs, N. Donaldson, S. P. Lacour, S. B. McMahon, and J. W. Fawcett, "A microchannel neuroprosthesis for bladder control after spinal cord injury in rat.," *Science translational medicine*, vol. 5, 2013.

[61] W. M. Grill and J. T. Mortimer, "Quantification of recruitment properties of multiple contact cuff electrodes.," *IEEE transactions on rehabilitation engineering*, vol. 4, pp. 49–62, June 1996.

[62] W. M. Grill and J. T. Mortimer, "Stability of the Input Output Properties of Nerve Cuff Stimulating Electrodes," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 4, pp. 364–373, 1998.

[63] J. Overgaard, L. Morten, M. Haugland, T. Sinkjaer, J. Larsen, and M. Thomsen, "Degeneration and regeneration in rabbit peripheral nerve with long-term nerve cuff electrode implant: a stereological study of myelinated and unmyelinated axons," *Acta neuropathologica*, pp. 365–378, 1998.

[64] C. Veraart, W. M. Grill, J. T. Mortimer, W. M. Grill, T. Mortimer, C. Veraart, and J. T. Mortimer, "Selective control of muscle activation with a multipolar nerve cuff electrode.," *IEEE transactions on bio-medical engineering*, vol. 40, pp. 640–53, July 1993.

[65] S. Jezernik, J. Wen, and N. Rijkhoff, "Whole nerve cuff recordings from nerves innervating the urinary bladder," *Second Annual IFESS*, 1997.

[66] R. R. Riso, F. K. Mosallaie, W. Jensen, and T. Sinkjaer, "Nerve cuff recordings of muscle afferent activity from tibial and peroneal nerves in rabbit during passive ankle motion.," *IEEE transactions on rehabilitation engineering*, vol. 8, pp. 244–58, June 2000.

Chapter 3

Velocity Selective Recording

3.1 Principles of Velocity Selective Recording

If it were possible to construct an ideal nerve interface, such a system would allow recording from and stimulation of every single neuron within the nerve; it would be stable in time so that each axon, once identified, would have a known function. Every axon would have a physiological label so, for example, some axons in the femoral nerve, after identification, would be labelled *vastus lateralis*, efferent.

However, before these physiological labels could be attached, it would be helpful to know the propagation velocity and direction for every fibre (afferent or efferent), which would greatly reduce its possible range of function. At present of course, no such interface exists. Practical methods start with the axon-specific, such as intrafascicular devices, fine tungsten needle electrodes [1] or arrays such as the Utah design [2]; these are typically invasive, show poor chronic applicability in peripheral nerves and give no indication of the axon or *action potential* (AP) characteristics. At the extrafascicular level, cuffs can be safe for chronic clinical use, but are limited in recording the composite activity from all the axons in the nerve. In this range of types, *longitudinal intrafascicular electrode* (LIFE) arrays can record from small groups of axons, perhaps within one fascicle, but do not show activity in the whole nerve [3]. Micro-channel nerve interfaces have enabled interfascicular recording from peripheral nerves with single unit activity resolution, but the number of axons in each micro-channel is quite large ($\approx$100 [4]) and physiological characterisation is still limited. No current interface method allows communication with one axon, or even a group of few axons, with physiological labelling.

The method of *velocity selective recording* (VSR) has been applied to cuffs and, by extension, could be applied to micro-channels. By filtering the neural signal in the *velocity domain*, activity within bands of conduction velocity can be discriminated and, if that band corresponds to a functional group of fibres, for example, the γ efferents, which are responsible for muscle spindle contraction, it should be possible to estimate the activity in those fibres. Thus the VSR method should improve both the quantity and quality of the

information that can be extracted from the neural signal using practical types of electrode structure. This improvement may yield substantial benefits in a clinical neuroprosthesis if one can provide better resolution at the input.

However, in spite of the many potential applications of VSR, until very recently it has only been demonstrated with electrically evoked *electroneurogram* (ENG), i.e. *compound action potentials* (CAPs) [5, 6]. This is because there are several very significant differences between the requirements of recording electrically-evoked and the natural ENG that complicate the process of recording the latter. These differences include much smaller signal amplitudes (using cuffs, typically 1 μV − 10 μV, as opposed to about 100 μV for CAPs) and the need to determine the rate of neural firing in a particular velocity band, rather than the relative amplitudes of activity between bands, which is generally the case for CAPs.

This chapter describes the existing theory of VSR as well as some of the limitations and practical considerations, an overview of alternative methods is also provided for purposes of comparison. The basic principle of VSR is to transform time domain recordings of the ENG into the velocity domain in a manner akin to the Fourier transform. The velocity spectrum, which is computed directly from the time domain recordings, represents both the direction of AP propagation and the level of excitation of each neuron within the nerve.

3.1.1 Delay-and-Add

The essence of VSR is a simple process called *delay-and-add* that is analogous to beamforming algorithms used in certain types of synthetic aperture arrays [7]. Channels of data recorded from linearly spaced electrodes along the length of the nerve are delayed relative to the last channel V_{B1} by an interval that depends on both the electrode spacing and the propagation velocity of the signal. So if the delay between the first two channels (V_{B1}, V_{B2}) is dt, the delay between the first and third channels (V_{B1}, V_{B3}) is $2.dt$ and so on. The general index of this process is i where $1 \leq i \leq C$ and where $C.dt$ is the maximum delay of interest. Delay-and-add operates by inserting delays into the channels to effectively cancel the naturally occurring delays, after which the channels' outputs are summed resulting in a single signal V_D (Equation 3.1.1) as illustrated in Figure 3.1.1 [5]. Inspection of Equation 3.1.1 reveals that the resultant velocity waveforms ($V_d[n, dt]$) are not linearly spaced as $dt = d/v$, that is to say that intrinsically there is greater resolution at the *slower* velocities. Of course it is possible to linearise the system by selecting discrete values of dt that correspond to linearly spaced velocities.

$$V_D[n, dt] = \sum_{i=1}^{C} V_{Bi}[n - (i-1).dt] \tag{3.1.1}$$

Where C is the number of recording channels and n is the current sample index.

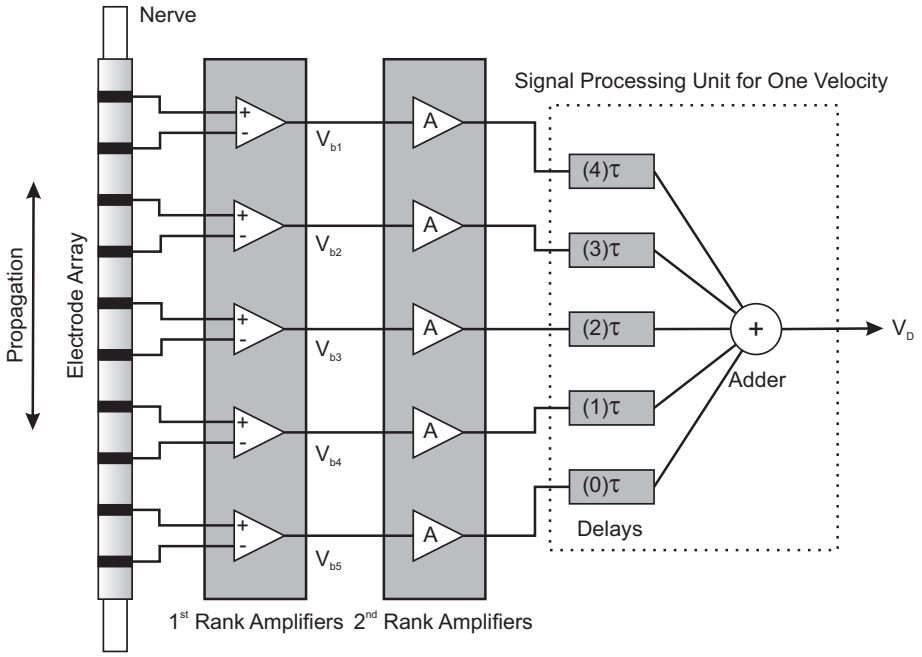


Figure 3.1.1: The basic concept of delay-and-add showing the signal processing requirements for one velocity. The signals are differentially amplified before being artificially delayed and summed together. The delay units required vary from one velocity to another.

For each velocity of interest, a corresponding value of dt is applied, so that there are many (m) data streams after the addition operation, each of which corresponds to a different conduction velocity. One advantage of VSR is the ability to record and isolate neural activity that is both *afferent* and *efferent* by simply applying a negative value of dt during the delay process. Furthermore, when delay-and-add is used and the noise sources are uncorrelated then there is an increase in *signal-to-noise-ratio* (SNR) of approximately $\sqrt{C}$ [8]. This property can be exploited to identify APs that may not be observable directly in the time records of the individual channels and thus could not be classified by traditional time domain based methods such as spike sorting.

3.1.2 Intrinsic Velocity Spectra

Various methods have been described in existing studies to estimate AP conduction velocity and to discriminate between *afferent* and *efferent* signals, the most basic of which employ a simple two electrode system [9]. Intuitively, artificially delaying the recorded signal from successive electrodes before summing them together gives a maximal response when the inserted delay is matched to the naturally occurring delay. The advantage to using multiple electrodes is two fold; firstly using multiple electrode channels provides an averaging effect that counteracts the uncorrelated noise sources, and secondly, the multiple electrodes may be connected in a tripolar configuration, which is effective at reducing interference from surrounding tissue [10, 11]. In order to identify the velocity profile of a recording, the maximum value of each $V_d[n, dt]$ is found and plotted against velocity; this

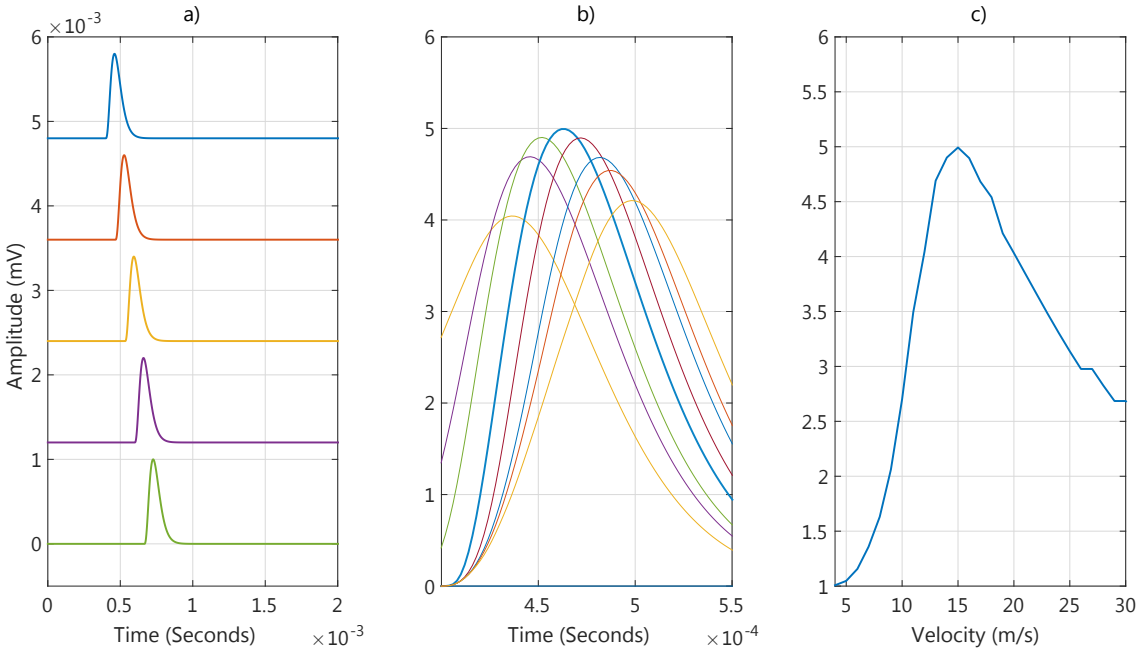


Figure 3.1.2: The process by which the IVS is generated from a segment of time domain data. **Inset a)** represents an ideal AP recorded with naturally occurring delays over five electrode channels, **inset b)** represent the V_D waveforms generated by the delay-and-add process, the time axis has been adjusted for the purposes of presentation. The largest response (or the *matched velocity*) is clearly visible. **Inset c)** is the IVS computed by taking the maximumn value of the V_D waveforms over the entire recordings.

profile is called the *intrinsic velocity spectrum* (IVS) and is a valuable tool for analysing the velocity content (spectrum) of a recording. The overall process is illustrated in Figure 3.1.2 that shows five channels of ideal data, the corresponding $V_d[n, dt]$ waveforms and the resulting IVS respectively.

3.1.3 Velocity Selectivity

Within a single fascicle there will be many individual neurons, each with independent axonal diameters and conduction velocities. Within the *sural* nerve of *man*, for example, there is a distribution of neurons ranging from 2 μm $-$ 12 μm in diameter (as shown in the histogram of Figure 3.1.3) [12]. It has already been shown that for myelinated axons there is an approximately linear relationship between axonal diameter and conduction velocity and within the sural nerve this relationship is approximately 5 m/s per μm, so the range of expected conduction velocities within the nerve would be about 10 m/s $-$ 60 m/s. Even when considering a single nerve, there is a wide range of different conduction velocities that need to be individually discriminated, in terms of both direction of propagation and conduction velocity. It is necessary to have a quantitative measure of how well APs with similar conduction velocities may be discriminated; this measure is called the *velocity selectivity*. The velocity selectivity is analogous to the *Q factor* in the frequency domain and is a measure of the broadness of the peak in the velocity domain. The velocity

selectivity, expressed as Q_v, is given by Equation 3.1.2:

$$Q_v = \frac{V_0}{V_{3+} - V_{3-}} \tag{3.1.2}$$

Where V_0 is the centre (or peak) velocity and V_{3+} and V_{3-} represent the velocities at which the amplitude has been reduced by a factor of 3 dB from the peak amplitude [8].

The velocity selectivity provides a useful metric by which improvements to the delay-and-add process may be evaluated and it is a useful tool for the quantitative analysis of a velocity selectivity based recording system. Although a fundamental limitation is imposed by the dependence of Q_v on the form of the AP in the time domain, and so when using the velocity selectivity to benchmark VSR systems it is important to maintain consistency in the APs that are used. In Figure 3.1.4 the velocity quality factor is illustrated for two simulated APs, the largest peak occurring at a conduction velocity of 35 m/s. Both the IVS and also an artificially enhanced spectrum are shown, the velocity quality factors are $Q_{ivs} = 1.6$ and $Q_{enhanced} = 11.3$ respectively. Within the intrinsic spectrum the large peak occurring at 35 m/s partially occludes the lower amplitude peak occurring at 41 m/s; if the simulated axons were closer in diameter then it is likely that the AP propagating at 41 m/s would be obscured entirely. In the enhanced spectrum however the improvement of the velocity selectivity by a factor of approximately 7 times permits a greater level of discrimination from one velocity to another.

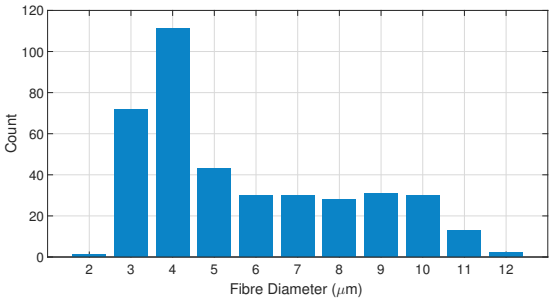


Figure 3.1.3: Histogram of the axonal diameters identified from histology within the sural nerve of man [12].

In the context of neural recording systems it is important to maximise Q_v for a given electrode structure so that the highest possible axonal resolution is possible. An example application for a closed loop neural prosthesis is the restoration of control of the urinary bladder following spinal cord injury. It has been shown in *man* that post-injury the sacral nerve roots that innervate the bladder (the pre-ganglionic sacral roots) continue to propagate APs that contain information about the fullness of the bladder [13]. A combined electrophysiological and histological study on these nerve fibres found two main distributions of neurons that encoded the bladder fullness, the single stretch $S1$ and the tension receptor ST afferents, with conduction velocities of 41 m/s and 35 m/s respectively. The peaks within the velocity spectra shown in Figure 3.1.4 may now be relabelled as the $S1$ and ST afferents, which individually encode the bladder volume and the bladder pressure, both of which are required in order to estimate the relative *fullness*. As the urinary bladder fills, the firing rates of the $S1$ and ST afferents vary from a baseline firing rate of 15 APs per 200 ms to a rate representing a full bladder of 40 APs per 200 ms [13]. As the bladder fullness is encoded in both the $S1$ and ST afferents it is

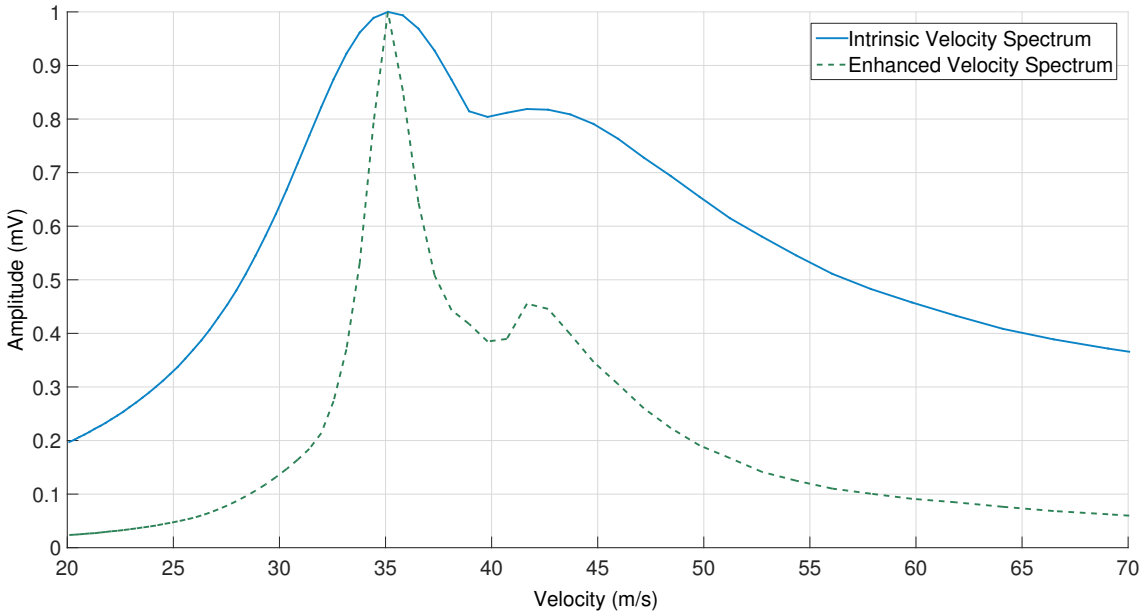


Figure 3.1.4: The IVS (solid line) and an enhanced velocity spectrum (dashed line) for two APs propagating at 35 m/s and 41 m/s respectively.

necessary to have a highly selective system, with $Q_v \geq 10$ in order to record accurately from the sacral roots.

3.1.4 Electrode Geometry and Amplifier Configurations

There are a number of different ways in which the electrodes within a *multiple electrode cuff* (MEC) may be connected to the recording amplifiers. A simple recording may be formed by the use of a single electrode located centrally within the recording assembly, a second remote electrode may be placed at some distance as a reference. This *monopolar* arrangement produces large signal amplitudes but the rejection of common mode interference is poor. Differential, or *bipolar* recording decreases the sensitivity of the electrode to external noise sources. In the bipolar configuration two electrodes within the recording assembly are connected directly to the inputs of a differential amplifier. The recorded signal amplitude now depends largely on the electrode separation as well as the inter-electrode impedance and the length of the recording assembly. A third approach is the *tripolar* configuration in which three amplifiers are used to produce a single recorded signal as shown in Figure 3.1.5. The two first-rank amplifiers record differentially between the outer electrodes, and the centre electrode acts as a reference. A third (second-rank) amplifier sums the output signals from the first-rank amplifiers.

The input to the MEC can be modelled as a *transmembrane action potential* (TMAP) denoted by $V_m(t)$ with frequency spectrum $V_m(f)$. The TMAP and its spectrum can be represented by the Fourier transform pair:

$$V_m(t) = At^n e^{-Bt} \tag{3.1.3}$$

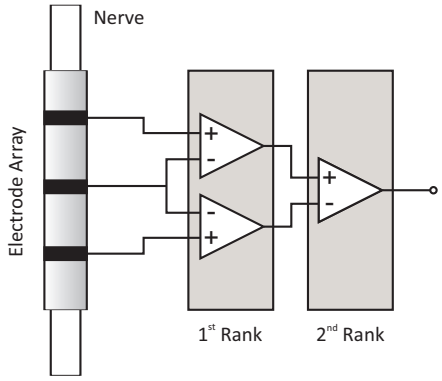


Figure 3.1.5: The true-tripolar amplifier configuration illustrated for a three electrode cuff.

$$V_m(f) = \frac{n!A}{(B + 2j\pi f)^{n+1}} \tag{3.1.4}$$

Where A and B are constants that define the shape of the TMAP and f is frequency [8]. Note that Equation 3.1.3 is only valid for $t > 0$.

In order to examine the effect of the electrode structure on the recorded signals it is necessary to consider a one-dimensional analytical model of the system. Figure 3.1.6 represents a single myelinated axon constricted by a narrow circumferential cuff of length L with a single recording electrode placed at position Z_0, this simplified model assumes an integer number of nodes of Ranvier, with nodes at either end of the cuff. Over the cuff length, the intra and extracellular resistances (r_i and r_e) are assumed to be uniform and equally divided into n inter-nodal regions by nodes $0, 1, 2, \ldots, n$, and so:

$$R_e = \sum_{i=1}^{n} r_e = nr_e \text{ and } R_i = \sum_{i=1}^{n} = nR_i \tag{3.1.5}$$

It is assumed that the resistance outside the confines of the cuff is negligible, and therefore the external end nodes are both grounded (nodes $V_{m(j-k)}$ and $V_{m(j+k)}$). The voltage generators $V_{m(i)}$, where $0 \leq i \leq n$, represent the potential differences across the membrane at the nodes of Ranvier, i.e the TMAP. At each node the TMAP is delayed by z/v, where v is the conduction velocity and z is the distance travelled. In order to determine the potential V_j recorded by a single electrode placed j nodes from the end of the cuff, in this case at location Z_0, disconnect point Q from ground and find the Thévenin equivalent voltage at Q. Starting with a single inter-node and taking a voltage loop:

$$-V_{m(i)} + V_{m(i)} = r_i I + r_e I \tag{3.1.6}$$

$$\therefore V_{th} = \frac{r_e}{r_i + r_e}\left[V_{m(i)} - V_{m(i-1)}\right] \tag{3.1.7}$$

Now considering all of the nodes:

$$V_Q = \frac{R_e}{R_i + R_e} \sum_{i=1}^{n} \left(V_{m(i)} - V_{m(i-1)}\right) \tag{3.1.8}$$

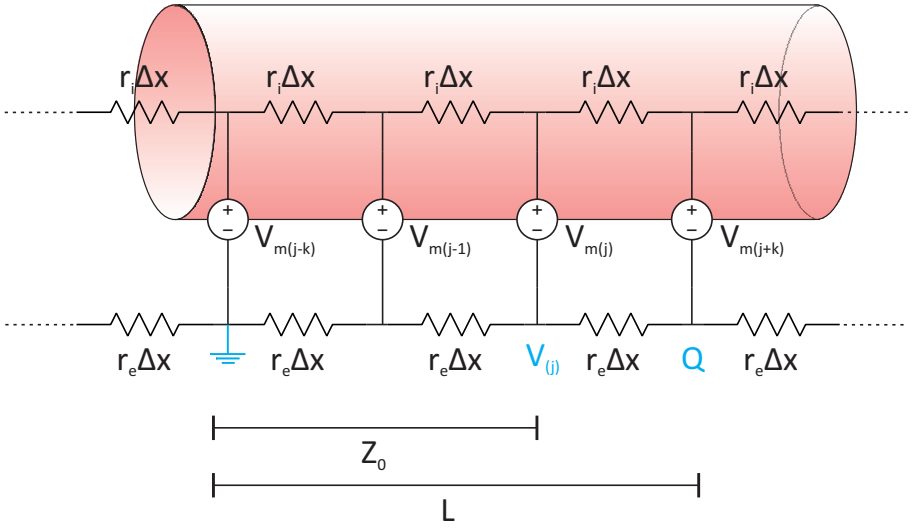


Figure 3.1.6: One-dimensional model used for the analytical analysis of a small cuff wrapped around a single axon, L is the length of the cuff and Z_0 is the location of a recording electrode. The nodes of Ranvier are modelled as voltage generators V_m and r_e and r_i are the inter-nodal intra and extracellular resistances respectively.

The equivalent resistance R_Q is found by shorting all voltage sources $V_{m(i)}$ and is:

$$R_Q = \frac{R_e R_i}{R_e + R_i} \tag{3.1.9}$$

When reconnected, a current of V_Q/R_Q flows out of the network at Q:

$$I_Q = \frac{V_Q}{R_Q} = \frac{1}{R_i} \sum_{i=1}^{n} \left(V_{m(i)} - V_{m(i-1)} \right) \tag{3.1.10}$$

Of I_Q, only a fraction will flow though the extracellular space:

$$I_Q \frac{R_i}{R_e + R_i} = \frac{1}{R_e + R_i} \sum_{i=1}^{n} \left(V_{m(i)} - V_{m(i-1)} \right) \tag{3.1.11}$$

The volt drop at point Z_0 can be found by considering the voltage developed by the fraction of R_Q that flows through the resistance R_e scaled by the ratio of Z_0/L:

$$V_Z = \frac{R_e}{R_e + R_i} \left[\frac{Z_o}{L} \right] \sum_{i=1}^{n} \left(V_{m(i)} - V_{m(i-1)} \right) \tag{3.1.12}$$

$$\therefore V_j = V_Q - V_Z = \frac{R_e}{R_e + R_i} \left[\sum_{i=1}^{n} \left(V_{m(i)} - V_{m(i-1)} \right) \frac{Z_0}{L} - \sum_{i=1}^{j} \left(V_{m(i)} - V_{m(i-1)} \right) \right] \tag{3.1.13}$$

The summation terms in Equation 3.1.13 may be simplified to give:

$$V_j = \frac{R_e}{R_e + R_i} \left[\left(V_{m(n)} - V_{m(0)} \right) \frac{Z_0}{L} - \left(V_{m(j)} - V_{m(0)} \right) \right] \tag{3.1.14}$$

Now, by observing that due to the delay introduced by the conduction velocity:

$$V_{m(0)} = V_m(t) \tag{3.1.15}$$

$$V_{m(j)} = V_m\left(t - \frac{Z_0}{V}\right) \tag{3.1.16}$$

$$V_{m(n)} = V_m\left(t - \frac{L}{V}\right) \tag{3.1.17}$$

Equation 3.1.14 may now be written in the form originally given by Stein and Pearson as [14]:

$$V_{mp}(t, Z_0) = \frac{R_e}{R_e + R_i}\left[\left(1 - \frac{Z_0}{L}\right)V_m(t) - V_m\left(t - \frac{Z_0}{v}\right) + \left(\frac{Z_0}{L}\right)V_m\left(t - \frac{L}{v}\right)\right] \tag{3.1.18}$$

Where $V_{mp}(t, Z_0)$ is the monopolar recorded signal, $V_m(t)$ is the TMAP, R_e and R_i are the extracellular and intracellular resistances and v is the conduction velocity of the AP.

Using this model V_m can be considered to be the input to the cuff and V_{mp} the output, thus it becomes possible to define a transfer function for the cuff. Equation 3.1.18 highlights the effect of the cuff end terms on the transfer function, i.e that the transfer function for a monopolar recording is not invariant of the location of the electrode within the cuff or of the length of the cuff. Extending the model to a three electrode system allows the use of a tripolar electrode configuration similar to that shown in Figure 3.1.5. The response of the tripole is shown in Equation 3.1.19 and the resulting cuff transfer function in Equation 3.1.20.

$$V_{tp} = V_{mp1}(t) - 2V_{mp2}(t) + V_{mp3}(t) \tag{3.1.19}$$

Where V_{mpi} is the signal recorded from each electrode within the cuff.

$$V_{tp}(t, Z_1, Z_2, Z_3) = \frac{R_e}{R_e + R_i}\left[V_m\left(t - \frac{Z_1}{v}\right) - 2V_m\left(t - \frac{Z_2}{v}\right) + V_m\left(t - \frac{Z_3}{v}\right)\right] \tag{3.1.20}$$

This is an important result as it shows that if the electrodes are equally spaced then a tripolar signal may be recorded from any point within the length of the cuff and it will be independent of potential gradients applied to either end of the cuff. This result reduces the effect of common mode signals from nearby tissues or muscle but it also permits the recording from many tripoles inserted within the length of the cuff independently from each other.

The tripole configuration has been shown to act as a spatial filter [15]. For the sake of analysis the TMAP (V_m) can be considered to be an impulse, or Dirac Delta function $\delta(t)$, then using Equation 3.1.20 the spatial impulse response of the symmetrical tripole can be estimated by Equation 3.1.21.

$$h(t) = \frac{R_e}{R_e + R_i}\left[\delta\left(t - \frac{Z_1}{v}\right) - 2\delta\left(t - \frac{Z_2}{v}\right) + \delta\left(t - \frac{Z_3}{v}\right)\right] \tag{3.1.21}$$

Where Z_1, Z_2 & Z_3 are the electrode positions within the cuff, and v is the conduction velocity of the axon.

Taking the discrete time Fourier transform of Equation 3.1.21 gives the frequency response shown in Equation 3.1.22.

$$H(\omega) = \frac{R_e}{R_e + R_i}\left[e^{-\frac{j\omega Z_1}{v}} - 2e^{-\frac{j\omega Z_2}{v}} + e^{-\frac{j\omega Z_3}{v}}\right] \tag{3.1.22}$$

Where $H(\omega)$ is the transfer function for a tripole and $\omega = 2\pi f$. Noting however that at matched velocities:

$$\frac{Z_2 - Z_1}{v} = \frac{Z_3 - Z_2}{v} = \frac{d}{v} = T \tag{3.1.23}$$

Equation 3.1.22 can be rearranged, starting by multiplying the Z_1 and Z_3 terms inside the brackets by $e^{-\frac{j\omega Z_2}{v}}$ to give:

$$H(\omega) = \frac{R_e}{R_e + R_i}\left[e^{-\frac{j\omega(Z_2 - Z_1)}{v}} - 2e^{-\frac{j\omega Z_2}{v}} + e^{-\frac{j\omega(Z_2 - Z_3)}{v}}\right] \tag{3.1.24}$$

Using Equation 3.1.23 and factoring out $e^{-\frac{j\omega Z_2}{v}}$ gives:

$$H(\omega) = \frac{R_e}{R_e + R_i}e^{-\frac{j\omega Z_2}{v}}\left[e^{j\omega T} - 2 + e^{-j\omega T}\right] \tag{3.1.25}$$

The terms inside the brackets may be factored to give:

$$H(\omega) = \frac{R_e}{R_e + R_i}e^{-\frac{j\omega Z_2}{v}}\left(e^{\frac{j\omega T}{2}} - e^{\frac{-j\omega T}{2}}\right)^2 \tag{3.1.26}$$

Then by using trigonometric substitution and neglecting the phase term, to give Equation 3.1.27.

$$|H(\omega)| = 4\frac{R_e}{R_e + R_i}sin^2\left(\frac{\omega T}{2}\right) \tag{3.1.27}$$

Combining $H(\omega)$ with $V_m(\omega)$ gives the overall frequency response of the cuff and tripole $Y(\omega)$ as shown in Equation 3.1.28 [16]. It is useful at this stage to note that $H(\omega)$ is a comb filter with a repeated passband of bandwidth $1/T$.

$$|Y(\omega)| = |H(\omega)||V_m(\omega)| = 4\frac{R_e}{R_e + R_i}sin^2\left(\frac{\omega T}{2}\right).\left|\frac{A}{(B + j\omega)^2}\right| \tag{3.1.28}$$

The transfer function of the delay-and-add stage may be found by noting that the phase response of $H(\omega)$ will vary for each tripole by a factor of $e^{-j\omega T}$ and that the artificial intertripole delay inserted by Equation 3.1.1 will produce an additional phase factor of $e^{j\omega\tau}$. The transfer function of the delay-and-add stage is therefore:

$$|G(\omega, \tau)| = \left|\sum_{i=1}^{N} e^{-j\omega(T-\tau)i}\right| \tag{3.1.29}$$

Evaluating the geometric progression in Equation 3.1.29 gives:

$$|G(\omega, \tau)| = \left|\frac{1 - e^{j\omega(T-\tau)N}}{1 - e^{j\omega(T-\tau)}}\right| \tag{3.1.30}$$

Which reduces to:

$$|G(\omega, \tau)| = \frac{sin(\omega(T - \tau)N)/2}{sin(\omega(T - \tau))/2} \tag{3.1.31}$$

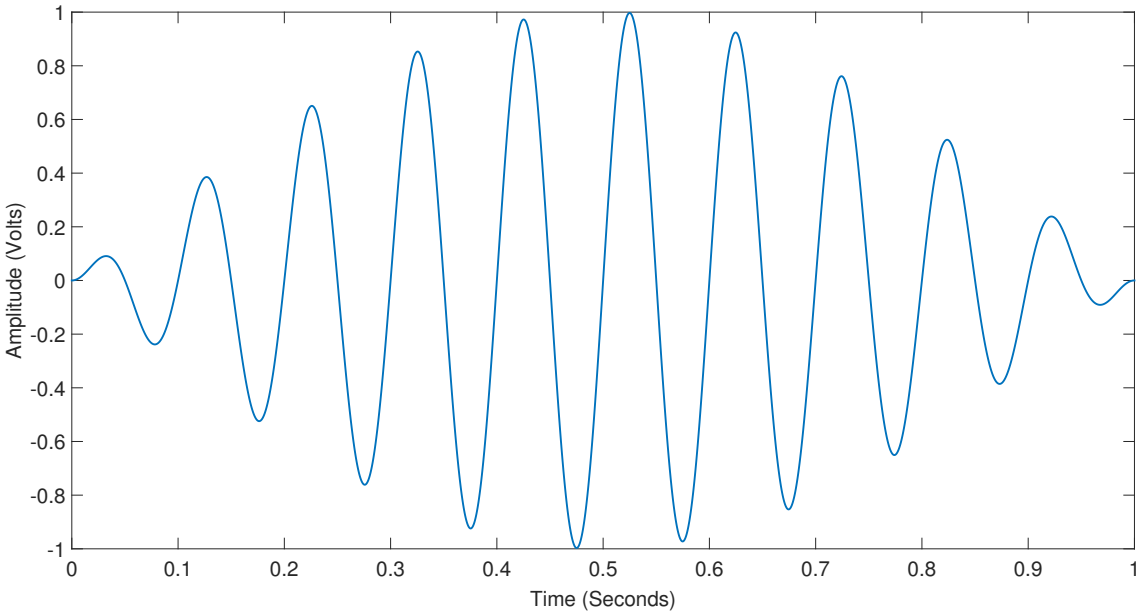


Figure 3.2.1: Time domain response of an 8^{th} order Butterworth band-pass filter of 20% relative bandwidth stimulated by a narrow pulse of unit amplitude.

If T and τ, the naturally occurring and artificially inserted delays are made equal, i.e the *matched* velocity then Equation 3.1.31 becomes:

$$|G(\omega,\tau)| = \lim_{\tau \to T} \frac{sin(\omega(T-\tau)N)/2}{sin(\omega(T-\tau))/2} \to N \tag{3.1.32}$$

Finally, the effect of the delay-and-add process may now be included into Equation 3.1.28, the overall transfer function $Y(\omega, v_0)$ is now the product of the spectrum of the TMAP ($V_m(\omega)$), the transfer function of one tripole ($H(\omega)$) and the transfer function of the delay-and-add process ($G(\omega)$ at a matched velocity $v = v_0$).

$$Y(\omega, v_0) = 4N\frac{R_e}{R_i} sin^2\left(\frac{\pi f d}{v}\right) . \left|\frac{An!}{(B+j2\pi f)^{n+1}}\right| \tag{3.1.33}$$

Equation 3.1.33 describes the output of a multiple electrode cuff with N tripoles, electrode spacing d and propagation velocity v and the matched velocity.

3.2 Methods for Increasing Velocity Selectivity

The intrinsic velocity selectivity is a limiting factor when designing a practical recording system based on VSR where there are constraints on the electrode geometry and the system sampling rate. A number of different approaches have been proposed in order to artificially increase the velocity selectivity; these methods will now be summarised.

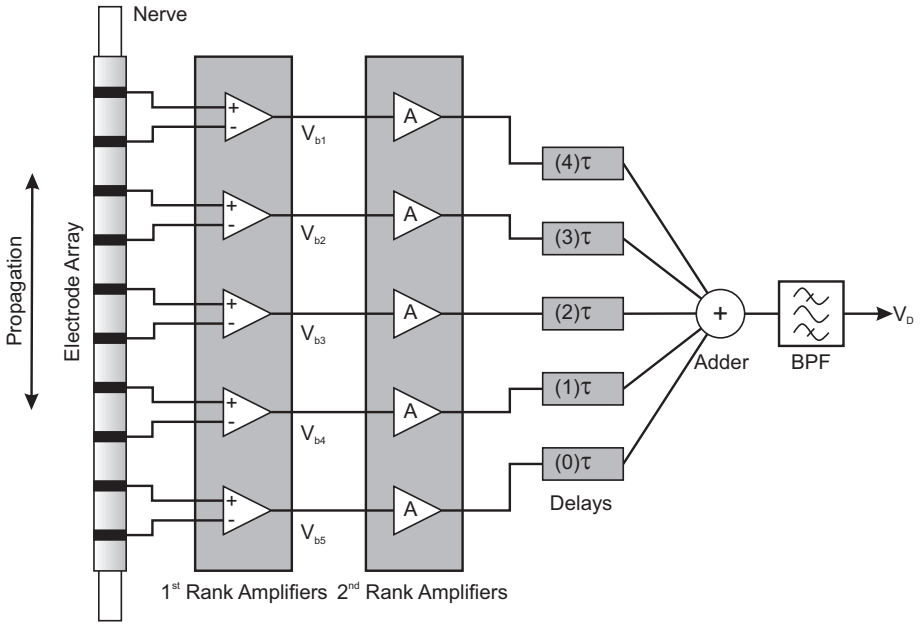


Figure 3.2.2: Modified VSR system with the addition of a band-pass filter after the summation component.

3.2.1 Band-pass Filter Structures

A proposed approach to increase the selectivity of the delay-and-add process is to place a band-pass filter at the output of each bipolar or tripolar amplifier [17, 18]. The impulse response of a band-pass filter is a burst of damped sine waves whose frequency is the centre frequency of the band-pass filter f_0. The duration of the burst as well as the amplitude of the burst depend largely on the order of the filter and the relative bandwidth. Figure 3.2.1 illustrates the impulse response of an 8^{th} order Butterworth band-pass filter of centre frequency 1 kHz and relative bandwidth 20%. The impulse applied to this filter was of single sample width and unit amplitude. To a good approximation, the TMAP as recorded by the amplifiers may be approximated to an impulse and so will produce similar bursts of sine waves when processed using a band-pass filter structure. The delay-and-add process is now effectively matching these bursts of sine waves as opposed to the TMAP itself, the addition of the band-pass filters in this way attempts to decouple the measurement of velocity from the spectral properties of the TMAP. Instead the velocity selectivity is now constrained by the parameters of the filter, which to some degree are free parameters. This proposed arrangement requires the use of a band-pass filter after each amplifier, in practice however the linearity of the system means that the order of the summation and filtering processes may be reversed. This leads to the much simpler and importantly more efficient architecture of Figure 3.2.2, where only a single band-pass filter is required per velocity.

In order to quantify the velocity selectivity available using band-pass filters the velocity

quality factor can be approximated as follows [17]:

$$Q_{bpf} = \frac{V_0}{V_{3+} - V_{3-}} \approx \frac{N\pi d}{2.64}\left(\frac{f_0}{v_0}\right) \tag{3.2.1}$$

Where N is the number of recording channels, d is the interelectrode spacing, f_0 is the centre frequency of the filter and v_0 is the matched velocity.

One of the limitations of using band-pass filter structures to enhance the performance of VSR is the potential for the generation of spurious velocities, or *images*. These occur due to the nature of the damped burst of sine waves; a single peak in the time domain is replaced with a burst of damped sine waves with frequency f_0 - the centre frequency of the band-pass filter. Therefore, although a global maximum occurs when all the channels are aligned perfectly, the potential exists for images to be generated when the channels are aligned by integer multiples of the period of the band-pass filter. Images will be generated if and only if Equation 3.2.2 is satisfied:

$$W \geq N \tag{3.2.2}$$

Where W is the number of periods of sine wave in each burst and N is the order of the system, or the number of channels. The velocity at which the images will occur is given by:

$$v_{image} = \frac{d}{\tau \pm q/f_0} \tag{3.2.3}$$

Where d is the interelectrode spacing and $\tau = d/v$ and q is an integer.

The velocities at which these images will appear is dependent on the centre frequency of the band-pass filter f_0, combined with the typical values expected in a practical recording system the images will tend to occur outside the conduction velocity band of interest (say 20 m/s – 120 m/s). It may therefore be possible to automatically reject such images from the desired signal by considering them to be *out of band*. Figure 3.2.3 illustrates the band-pass filtered velocity spectra for a single TMAP propagating at 30 m/s for six different centre frequencies. As the centre frequency is increased from 1 kHz to 32 kHz the velocity selectivity at 30 m/s increases from 0.8 to 33.2 respectively. However this increase in selectivity is accompanied by images within the velocity band starting at 7 m/s for a centre frequency of 2 kHz and increasing to 28 m/s for a centre frequency of 32 kHz. At lower centre frequencies the image could be considered to be out of band and rejected, but at higher centre frequencies the image becomes in-band and would be difficult to remove automatically. A second effect is that the energy in each spectrum is variable and tends to decrease monotonically with centre frequency (this is due to the form of the intrinsic spectrum of the TMAP) and so the effective peak-to-noise ratio is reduced.

Theoretical Limits

It is useful to understand the fundamental limits on the velocity selectivity that may be obtained via the use of band-pass filters. In order to determine analytically the upper and

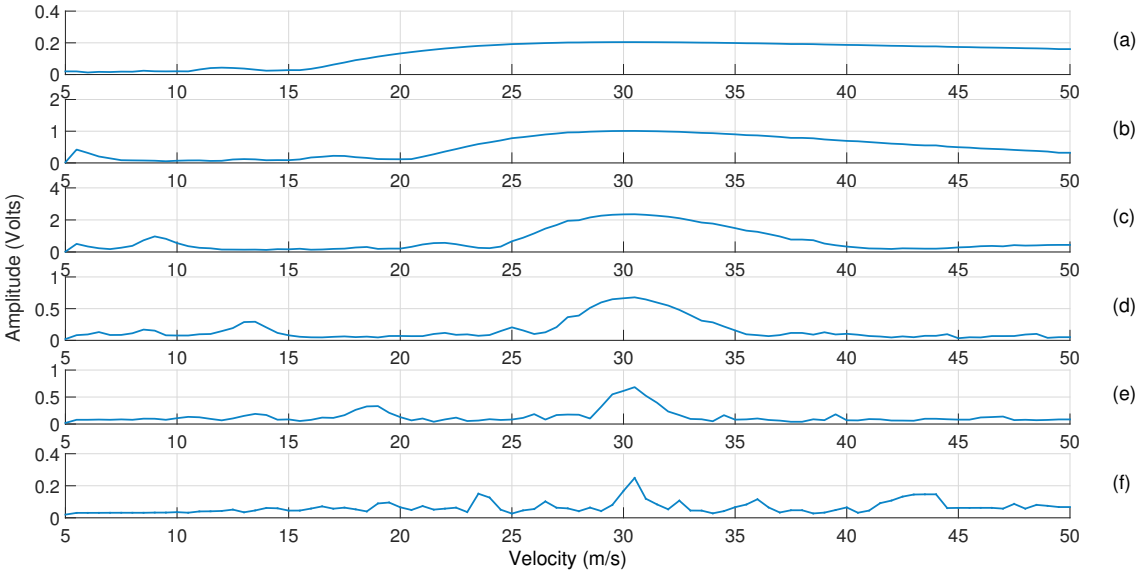


Figure 3.2.3: Band-pass filtered velocity spectra for a single TMAP propagating at 30 m/s. Filter centre frequencies and corresponding selectivitys are: a) 1 kHz - 0.8 b) 2 kHz - 2.03 c) 4 kHz - 4.06 d) 8 kHz - 6.7 e) 16 kHz - 15.2 f) 32 kHz - 33.2.

lower bounds on the velocity selectivity generated by a multiple electrode based recording system, we make the fundamental assumption that band-pass filters are always used and hence Equation 3.1.32 reduces to $Y(v)$, f appearing as a constant factor (f_0).

Using Equation 3.2.1, calculation of the range of available velocity selectivity (expressed as Q_v) at a velocity v_0 reduces to finding the range of possible values of f_0, which might appropriately be called the *selectivity bandwidth*. The lower bound of the selectivity bandwidth is taken to be the frequency of a sine wave whose width most closely matches the positive phase of an AP of the same amplitude, as illustrated in Figure 3.2.4. To a first order this is an approximation of the intrinsic velocity selectivity obtained without the use of band-pass filters. The alignment of the AP and the positive peak of the sine wave is carried out at the -3 dB points of both waveforms (the points at which the respective amplitudes have reduced by 3 dB of the peak values. At this point it can be shown that, to a good approximation, the equivalent frequency of the sine wave f_L is given by:

$$f_L \approx \frac{B}{8} \tag{3.2.4}$$

Where the parameter B relates to the time domain model of the AP ($V_m(t)$) of Equation 3.1.3. For example, if $B = 15$ kHz then $f_L = 1.875$ kHz and hence substituting into Equation 3.2.1 with $N = 10$ electrodes and an interelectrode spacing of $d = 3$ mm gives a lower limit (or the intrinsic value) of the velocity selectivity at 30 m/s of $Q_v = 2.2$.

It has been shown that the upper bound on the velocity selectivity is constrained by noise considerations because the spectrum of the signal (see Equation 3.1.4) decreases monotonically with frequency and eventually becomes indistinguishable from the noise floor of the system [17]. Beyond this interception with the noise floor it is clearly impossible

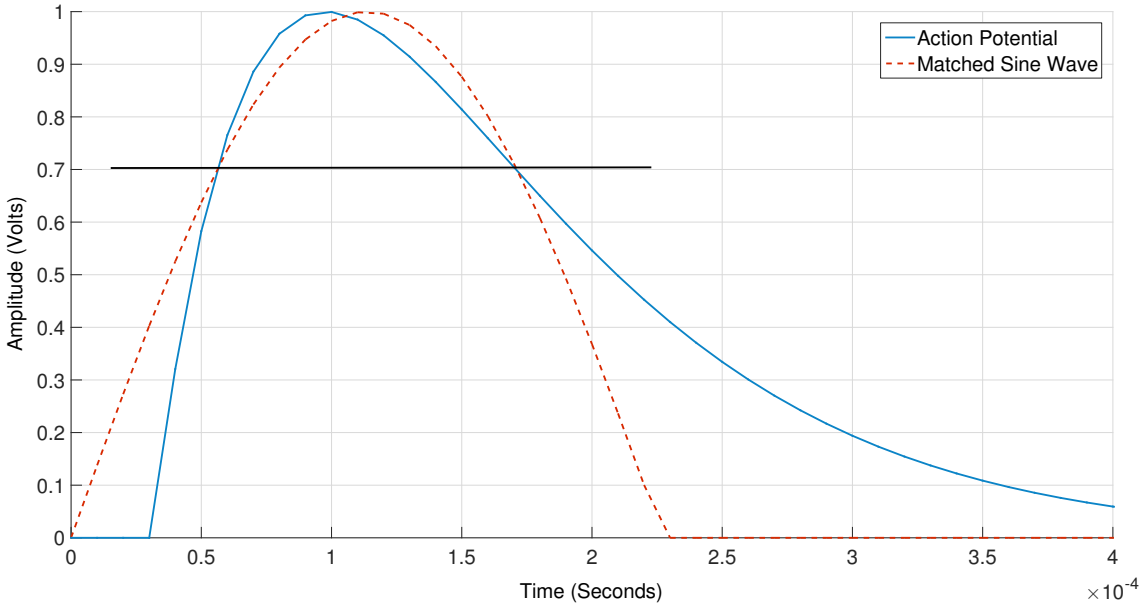


Figure 3.2.4: Fitting half a sine wave (dashed line) to an AP (solid line) to find the lower bound on velocity selectivity. The matching is in the time domain at the −3 dB points of the curves.

to distinguish the signal from the noise and so it seems reasonable to choose this *noise corner frequency* as the upper frequency limit for a band-pass filter, this in turn will limit the maximum available velocity selectivity. In the example shown in Figure 3.2.5 the spectrum of a single monopolar TMAP is plotted both with and without *additive white Gaussian noise* (AWGN). The AP is derived from Equation 3.1.3 with parameters $A = 40,774$ V, $B = 15$ kHz and $n = 1$. These constants were chosen to be representative of a mammalian AP from a myelinated axon normalised to a peak amplitude of unity. The noise generator employed produces zero mean white Gaussian noise with instantaneous power σ^2; so for a recording with length L the mean noise power is σ^2/L and the *root mean squared* (RMS) noise power is $\sigma/\sqrt{L}$. For example if $\sigma = 0.1$ and $L = 1024$ then the RMS noise power is 3.125×10^{-3} W. In Figure 3.2.5 the spectrum of the noise is not flat as would be expected for white Gaussian noise, this is because the Fourier transformation is not particularly effective when performing spectral power density estimation on random signals, a more accurate representation could be achieved by using a periodogram or a correlogram. In this simple case however the underlying characteristics of the noise generator were well understood and so the FFT was used to provide only an illustrative comparison.

The analytical form of the spectrum of the TMAP is given in Equation 3.1.3 and was found using the continuous time Fourier transform, unlike the spectrum plotted in Figure 3.2.5 that was calculated using a 1024 point fast Fourier transformation with a sampling frequency of 100 kHz. For frequencies well below the Nyquist limit (50 kHz) the two representations for the spectrum are very similar, for frequencies closer to the Nyquist limit the two representations diverge somewhat. This was compensated for by using a

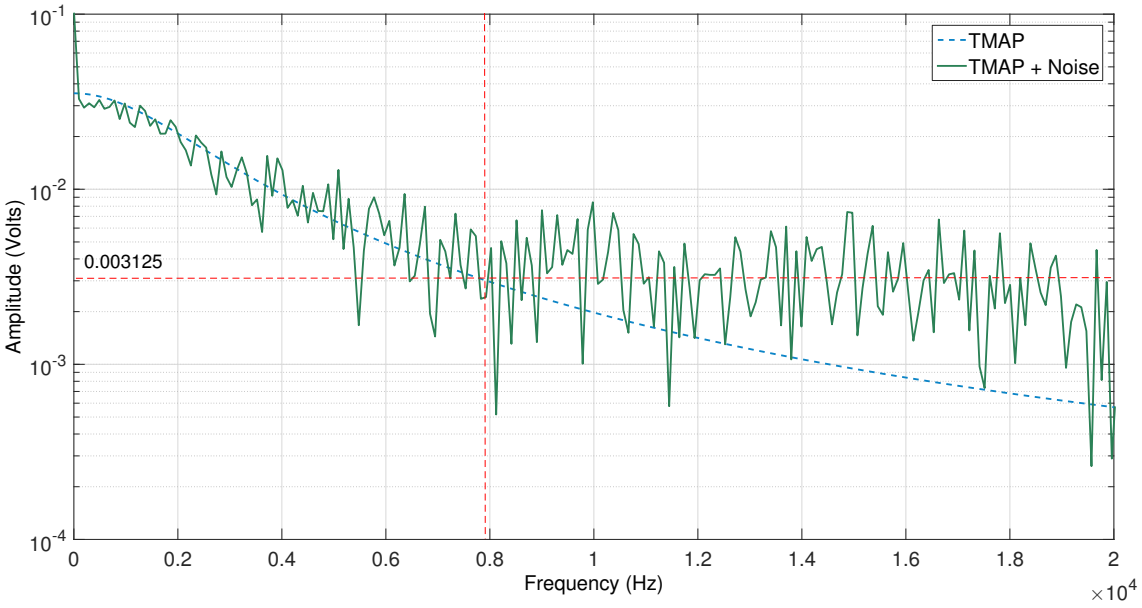


Figure 3.2.5: Spectrum (1024 point FFT) of a single monopolar AP (dashed curve) and the same with AWGN (solid curve). In this example $\sigma = 0.1$ ($SNR \approx 0$ dB) and so the noise floor is 3.125×10^{-3} V. The spectra intersect at a frequency of approximately 7.7 kHz. This *noise corner frequency* is taken to be the maximum frequency at which a band-pass filter can operate and therefore defines the maximum available velocity selectivity.

suitably high sampling rate, ensuring that the two representations were equivalent for the in-band frequencies.

Equation 3.2.5 is the magnitude of the spectrum of the TMAP, in order to calculate the noise corner frequency (the limit at which the energy in the signal is less than the energy in in the noise) ω_{lim} it is necessary to set $|V_m(\omega)|$ equal to σ/L and solve for ω_{lim}, resulting in Equation 3.2.6.

$$|V_m(\omega)| = \frac{2F_s}{L}\frac{A}{B^2+\omega^2} \tag{3.2.5}$$

$$\omega_{lim} = 2\pi f_{lim} = B\sqrt{\frac{2AF_s}{\sigma\sqrt{L}B^2} - 1} \tag{3.2.6}$$

It is now possible to combine Equations 3.2.5 and 3.2.6 to complete the expression for the selectivity bandwidth, given by Equation 3.2.7.

$$\frac{B}{8} \leq F_0 \leq \frac{B}{2\pi}\sqrt{\frac{2AF_s}{\sigma\sqrt{L}B^2} - 1} \tag{3.2.7}$$

It is beneficial to express the selectivity bandwidth in terms of the overall SNR. The average signal power is calculated from the time domain TMAP function:

$$P_s = A\sqrt{\frac{1}{\tau}\int_0^\tau t^2 e^{-2Bt}dt} \approx \frac{A}{2\sqrt{\tau}B^{3/2}} \tag{3.2.8}$$

Where τ is the length of the time domain recording and is given by $\tau = LT_s = L/F_s$.

Table 3.2.1: Relationship between SNR and the selectivity bandwidth.

Signal-To-Noise Ratio		Upper Limit on f_0 (kHz)		Selectivity Bandwidth (kHz)	
σ	**SNR**	**N = 1**	**N = 10**	**N= 1**	**N = 10**
0.01	10	25.3	45.16	23.3	43.16
0.1	0	7.7	14.12	5.7	12.12
1.0	-10	0.88	3.85	-	1.85

Combining this with the RMS noise power gives an expression for the overall SNR:

$$SNR = \frac{A\sqrt{F_s}}{2\sigma\sqrt{L}B^{3/2}} \tag{3.2.9}$$

Combining Equations 3.2.7 and 3.2.9, the selectivity bandwidth can be rewritten in a more compact form:

$$\frac{B}{8} \leq f_o \leq \frac{B}{2\pi}\sqrt{4\left(\frac{F_s}{B}\right)^{0.5}(SNR) - 1} \tag{3.2.10}$$

Recalling that the SNR increases as a function of $\sqrt{N}$ (where N is the number of channels) the final expression for the selectivity bandwidth can be found:

$$\frac{B}{8} \leq f_o \leq \frac{B}{2\pi}\sqrt{4\left(\frac{NF_s}{B}\right)^{0.5}(SNR) - 1} \tag{3.2.11}$$

Since the ratio of velocity selectivity enhancement, R_s, is proportional to this selectivity bandwidth (see Equation 3.2.1), it can be written as:

$$R_s \approx \frac{8}{\pi}\left(\frac{NF_s}{B}\right)^{0.25}(SNR)^{0.5} \tag{3.2.12}$$

For example, with the AP and system parameters discussed previously and a SNR of 0 dB, the ratio of velocity selectivity enhancement R_s is 11.7; which is a significant level of enhancement. Inspection of Equation 3.2.12 reveals that R_s changes very slowly as a function of N, F_s and B but varies much more rapidly with SNR. This result suggests that in a practical system design it may be more worthwhile to improve the intrinsic SNR through the use of more advanced amplifiers or electrodes than it would be to simply add more electrodes or to increase the sample rate.

Simulation Results

The analysis of the selectivity bandwidth was verified by simulation, a ten channel system was simulated using the parameters discussed above for four different levels of SNR ranging from ∞ dB to -10 dB. The resulting selectivity parameters are given in Table 3.2.1. Figure 3.2.6 shows the IVS for a single AP propagating at 30 m/s for the four different SNRs. For the case where the SNR is 10 dB the spectra is identical to the noise free case, as the SNR is decreased there is some degradation in performance. When the SNR reaches -10 dB there is no longer a distinguishable peak within the spectra and the method has failed. This is entirely in agreement with the values predicted in Table 3.2.1, the delay-and-add process fails when f_{lim} falls below the limit governed by Equation 3.2.6.

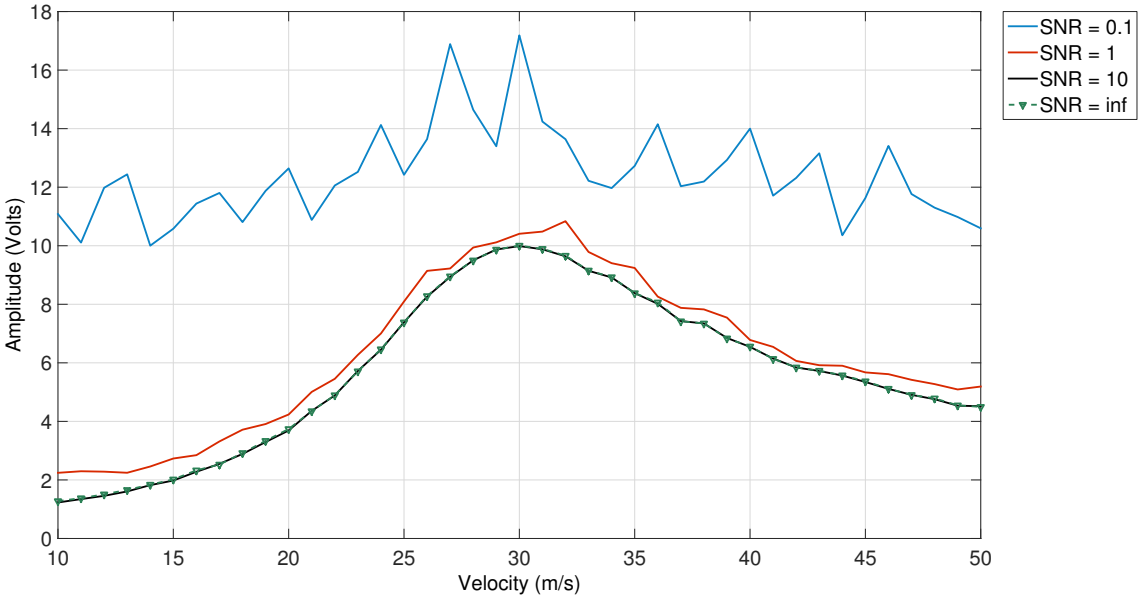


Figure 3.2.6: IVS for a 10 channel system with added AWGN. Three values of SNR are indicated (10, 1 and 0.1) in addition to the noiseless case. Note that for SNR = 10 the profile is indistinguishable from the noiseless case. For SNR = 1 there is some degradation of performance but for SNR = 0.1 the method fails completely.

3.2.2 Artificial Time Delay Neural Networks

Recently a new method for performing VSR of the ENG has been proposed that makes use of nonlinear artificial neural networks with added time delay components [19]. The artificial neural network employed is a modified feedforward type known as a *time delay neural network*, the modifications consist of a fixed length *finite impulse response* (FIR) filter added to each input of every artificial neuron [20]. This type of network has a history of applications such as speech recognition and time series prediction [21, 22]. As is the case with the delay-and-add structure, an entire (separate) neural network is required per velocity; Figure 3.2.7 provides an overview of the network structure. The inputs to the first layer are the normalised outputs from the amplifiers and are connected to a set of FIR filters realised as a bank of digital shift registers with programmable weights. The length of the shift registers is determined by the conduction velocity of interest and is a function both of the electrode spacing and the sampling rate (high sampling rates will necessitate longer shift registers). The outputs from the filters are then summed together with an additional bias parameter before being multiplied by a nonlinear element such as the logistic sigmoid function or the hyperbolic tangent.

One of the difficulties with time delay neural networks is the non causal nature of the instantaneous error; this prevents the use of traditional gradient descent-based training methods such as backpropagation [20]. Instead the network must be trained using random search methods that are significantly slower and are unlikely to find the global minima of the search space. Once trained however the network can be run effectively in real time

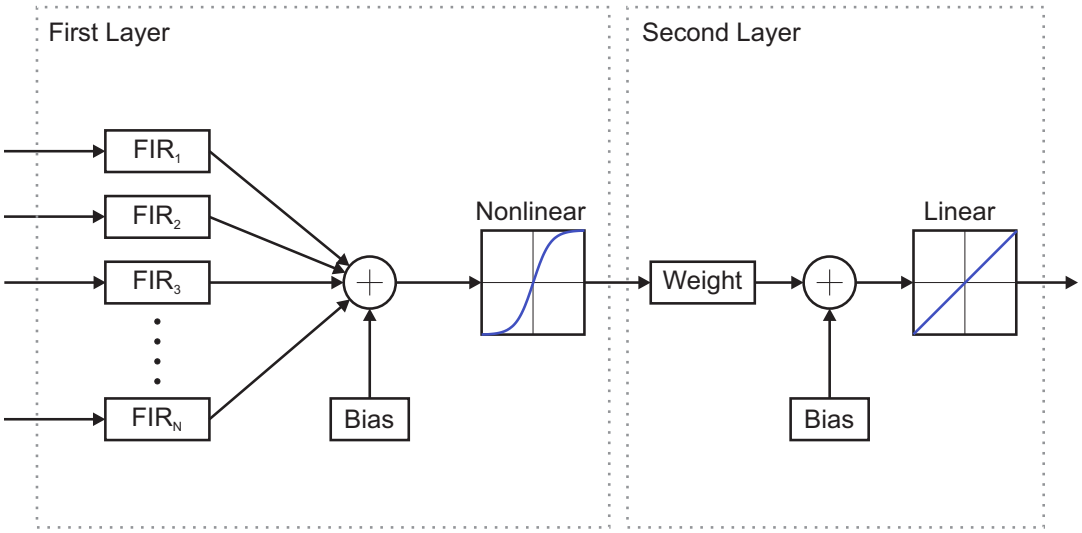


Figure 3.2.7: Modified artificial neural network consisting of two layers, the first of which contains a set of tapped delay lines with adaptive weights, the second layer features a linear transfer function.

Table 3.2.2: Comparative device utilization summary and power consumption [19].

Method	Parameter	Result
Delay-and-add	Total Equivalent Gate Count	867,187
	Current Consumption	19 mA
	Total Power Consumption (Measured)	22.8 mW
	Total Power Consumption (Estimated)	22.23 mW
TDNN	Total Equivalent Gate Count	1,563,624
	Current Consumption	30 mA
	Total Power Consumption (Measured)	45.6 mW
	Total Power Consumption (Estimated)	48.94 mW

and without any further supervision.

In a practical realisation the individual filter banks may be time division multiplexed in order to save hardware. Table 3.2.2 shows a comparative device utilization for the delay-and-add and time delay neural network approaches. The neural network requires approximately double the number of gates and double the power consumption for a single velocity. This increase however should be considered within the context that the designs were implemented using a *field programmable gate array* (FPGA), which is well-known to have significantly higher power consumption than an equivalent full custom device. The authors proposed that the power differences could be mitigated by implementing the neural network within a full custom 0.18 μm device [19]. An offline training process could be used to generate the necessary weights for each velocity and these could be stored within a memory structure on-chip and programmed into the neural network depending on the desired conduction velocity.

Simulated results using this method have shown increases in velocity selectivity of over ten times that achievable with the intrinsic velocity spectra however the authors did not examine the filter coefficients post training and so it is difficult to describe the exact mechanism by which this method operates. It is likely that the large gains in selectivity are partially attributable to the nonlinear element in the final layer; this element, if correctly biased, would be effective at compressing mismatched velocities and expanding matched velocities. When applied to data measured *in vivo* it was found that the neural network became highly unstable and frequently generates spurious responses. It is likely that the training process produces a highly optimised but also highly sensitive network that is easily disturbed by low frequency variations in the input waveforms. It may be possible to modify the network structure using an adaptive gain stage but at present the instability and sensitivity of this method have limited it to the analysis of simulated data *only*.

3.2.3 Beamforming Structures

Karimi et al. made the observation that the conventional delay-and-add process is equivalent to the beamforming techniques used in phased array radar signal processing [23]. Space-time beamforming is a commonly used technique for spatial filtering of broad-band signals in phased array antennas.

In this method FIR filters are inserted after the delay elements but before the summation element. The effect of the filters is to create a uniform frequency-angle beam pattern; and the same concept can be applied to delay-and-add. Figure 3.2.8 shows a modified delay-and-add structure with the extra FIR filter banks required in this method. The aim of these filters is to linearise the sensitivity of the delay-and-add process to different conduction velocities so that the masking effects detailed in Section 3.1.3 are minimised.

The weights for the filter banks must be derived numerically using minimisation approaches, however the weights obtained are generally complex numbers and so a quadrature filter (*Hilbert transformation*) must be inserted within the filter banks in order to process the real numbered samples. The basic structure is similar to that proposed for the time delay neural network except with an extra delay element, simulated results using this structure are promising although the authors did not perform a direct comparison to existing methods and only simulated results are available. Evidence suggests that it may be possible to approximately double the selectivity achievable using conventional band-pass filters if this method is employed, although the extensive training procedure may prove prohibitive and there is as yet no validation with recorded data.

3.2.4 Spatiotemporal Filtering

Recent developments by Soulier et al. have proposed a modified VSR system that uses spatiotemporal filtering in place of the delay-and-add structure [24]. Spectral analysis can

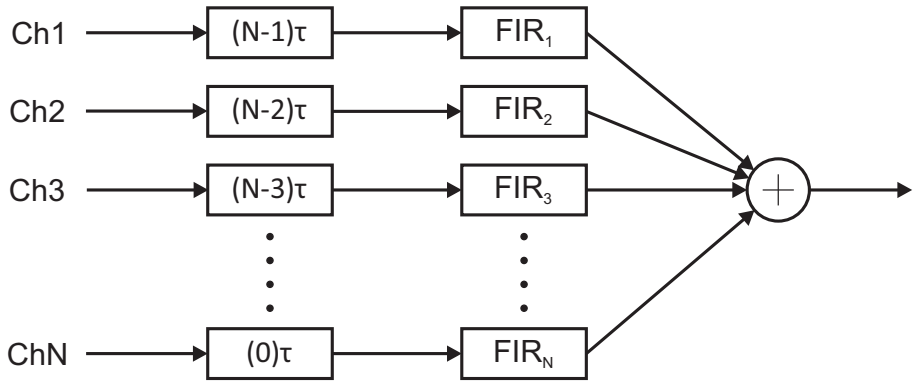


Figure 3.2.8: Modified delay-and-add structure incorporating FIR filters typical of space-time beamforming techniques.

be carried out in the temporal domain, the spatial domain or in both simultaneously. In the direction of the propagation of the signal, the spatial frequency domain f_z is related to the temporal frequency domain f_t by the conduction velocity:

$$v = \frac{f_t}{f_z} = \frac{d}{t} \tag{3.2.13}$$

Spatiotemporal filtering can be achieved by considering the individual amplifier channels as a quantised spatial domain, so that a two dimensional space may be formed that comprises of N amplifier channels sampled for M samples. Then the two dimensional discrete Fourier transform may be used to perform spectral analysis and filtering, essentially treating the M by N feature space as an image:

$$S(f_z, f_t) = \frac{1}{MN} \sum_{z=1}^{N} \sum_{t=1}^{M} s(z,t) e^{-i2\pi\left(\frac{f_s z}{N} + \frac{f_t t}{M}\right)} \tag{3.2.14}$$

Where $s(z,t)$ is the spatiotemporal signal and $S(f_z, f_t)$ is the signal in the spatial and temporal frequency domains.

Once the data has been transformed into the frequency domains, a filtering process (convolution) is applied using a simple filter $H(f_z, f_t)$ on the two-dimensional spectrum. The inverse Fourier transform can then be used to recover the filtered time domain waveforms:

$$v(f_z, f_t) = \sum_{f_z=1}^{N} \sum_{f_t=1}^{M} H(f_z, f_t) S(f_z, f_t) e^{i2\pi\left(\frac{f_s z}{N} + \frac{f_t t}{M}\right)} \tag{3.2.15}$$

Where $v(z,t)$ is the ideally filtered output signal and $H(f_z, f_t)S(f_z, f_t)$ is the masked data from the two-dimensional spectrum.

This method has shown promising results on simulated data, although only a small range of different conduction velocities were considered and the simulations were noiseless. It is not clear how noise will affect the selectivity of this method although it may be possible to combine this method with delay-and-add to exploit the effective increase in SNR. The signal processing required (essentially just the Fourier transform) is relatively simple and

this method may have considerable merit for a practical recording system if it can be demonstrated on biological recordings and in the presence of noise.

3.3 Other Methods for Action Potential Classification

3.3.1 The Spike Sorting Process

Extracellular recording of physiological neural activity is the dominant experimental technique in electrophysiology experiments, providing valuable information about both the central and peripheral nervous systems [25]. Extracellular electrodes such as hooks or suction electrodes are considered to be both reliable and minimally invasive recording approaches. These types of electrodes spontaneously record the electrical activity from an unknown number of excited axons which may serve different functions such as movement or memory. The study of these high-order functions often requires simultaneous recording from large areas of interconnected neurons using, for example, multi-channel electrode arrays [26]. There is a special interest in reconstructing the waveform of individual neurons from these multi-channel recordings so that the firing rates, or *spike trains*, may be extracted and compared. This procedure is referred to as neural *spike sorting*: the labelling and classification of individual APs based on morphology. It is generally considered that the shape and amplitude of APs recorded from a single neuron are time invariant and are a function of both the axon diameter and the distance from the neuron to the recording electrode.

Historically the spike sorting process was performed visually by a researcher who, after first determining the number of different classes, would classify each AP based on shape and reconstruct individual spike trains for each neuron. This time consuming process was rapidly replaced by more reliable computer based automated methods that make use of modern statistical analysis and signal processing [27–29].

There are many different approaches to automated spike sorting and a more detailed analysis is provided in Chapter 9, but the majority can be broken down into seven main steps [25]:

Filtering - In the first stage the raw data recorded from a single extracellular electrode is band-pass filtered to remove the frequency components of the noise that are out of band. The frequency range of interest is typically 10 Hz $-$ 10,000 Hz.

Spike Detection - Individual APs are now identified using two main steps: the pre-emphasis of the signal and the use of an amplitude threshold. The threshold is generally determined automatically using an unsupervised training process [30]. More complex methods use advanced detectors such as the *nonlinear energy operator* (NEO) which attempts to locate rapid changes in energy [31].

Alignment - Once APs have been detected a window of data is extracted and the waveform is captured. At this point each AP is aligned based on a measure such as the maximum value or the point of maximal slope [32].

Feature Extraction - Often very simple features such as the maximum amplitude are used to describe the morphology of the AP [32]. More complex methods employ statistical concepts such as *principal component analysis* (PCA) to extract unique features from the captured APs [33].

Dimensionality Reduction - The resulting features have many dimensions that make identification computationally expensive; one method to reduce the dimensionality of the features is simply to subsample at regular intervals. There are more advanced statistical methods, such as *Hartigans dip test* [34]. When using PCA it is common to reduce the number of dimensions considered for clustering by simply truncating to N dimensions, where N is typically 2 or 3 [35].

Clustering - The reduced data set can now be automatically examined for clusters that correspond to different classes. Clustering (especially unsupervised) is often the most difficult part of the sorting process. The *de-facto* benchmark is k-means clustering, however this relies directly on human intervention and so is not suitable for unsupervised learning. One algorithm that is suitable for unsupervised clustering is *valley* seeking [36], although the algorithm is not real time and has serious drawbacks in terms of complexity. When dealing with clusters that are irregular in shape, improvements may be obtained with superparamagnetic clustering [25].

Classification - As new spikes are observed in the recorded data, the clusters are updated and can be used to classify new spikes. From this a visual record called a *spike train* can be produced that shows the firing patterns of each class of spike.

A secondary motivator for identifying neural activity to within functional groups is a reduction in the amount of raw data that is produced. Systems have been reported with over one hundred electrode channels and sampling rates of 30 kS/s; for a resolution of 10 bits such a system produces data at a rate of 30 Mb/s [37]. The strict power limits placed on implantable systems makes transmission of data at this rate impractical. Therefore some mechanism is required that can sort the raw data locally, before transmitting only the desired information which is encoded within that data.

There are two major drawbacks with any method that relies upon spike detection and classification. Firstly they are often unsuitable for real time processing due to a large computational overhead, secondly at multiple stages they require prior information about the APs in order to form the statistical models required for clustering and threshold calculation. These problems place limitations on the practical applications; high computational complexity limits processing to a clinical environment and precludes chronic implantation. The requirement for prior knowledge reduces the flexibility of the system to deal with variations within the AP and implies long periods of training and adjustment.

VSR has the potential to solve these problems in two important ways; primarily the computational overheads are minimal and the processing can easily be performed in real time, and there is no requirement for any prior knowledge more advanced than the already well understood classifications of conduction velocity.

References

[1] K. Yoshida, K. Jovanović, and R. B. Stein, "Intrafascicular electrodes for stimulation and recording from mudpuppy spinal roots," *Journal of Neuroscience Methods*, vol. 96, pp. 47–55, Mar. 2000.

[2] E. M. Maynard, C. T. Nordhausen, and R. A. Normann, "The Utah Intracortical Electrode Array: A recording structure for potential brain-computer interfaces," *Electroencephalography and Clinical Neurophysiology*, vol. 102, pp. 228–239, Mar. 1997.

[3] T. Boretius, J. Badia, A. Pascual-Font, M. Schuettler, X. Navarro, K. Yoshida, and T. Stieglitz, "A transverse intrafascicular multichannel electrode (TIME) to interface with the peripheral nerve.," *Biosensors & bioelectronics*, vol. 26, pp. 62–9, Sept. 2010.

[4] J. J. FitzGerald, N. Lago, S. Benmerah, J. Serra, C. P. Watling, R. E. Cameron, E. Tarte, S. P. Lacour, S. B. McMahon, and J. W. Fawcett, "A regenerative microchannel neural interface for recording from and stimulating peripheral axons in vivo.," *Journal of neural engineering*, vol. 9, Feb. 2012.

[5] M. Schuettler, V. Seetohul, N. J. M. Rijkhoff, F. V. Moeller, N. Donaldson, and J. Taylor, "Fibre-selective recording from peripheral nerves using a multiple-contact cuff: Report on pilot pig experiments.," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 3103–6, Aug. 2011.

[6] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[7] Y. Huang and J. P. Miller, "Phased-array processing for spike discrimination.," *Journal of neurophysiology*, vol. 92, pp. 1944–57, Sept. 2004.

[8] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

[9] J. Hoffer, "Techniques to study spinal-cord, peripheral nerve, and muscle activity in freely moving animals," *Neurophysiological Techniques*, 1991.

[10] A. Demosthenous and J. Taylor, "Design of an adaptive interference reduction system for nerve-cuff electrode recording," *IEEE Transactions on Circuits and Systems*, vol. 51, no. 4, pp. 629–639, 2004.

[11] R. Rieger and J. Taylor, "Design of a low-noise preamplifier for nerve cuff electrode recording," *IEEE Transactions on Solid-State Circuits*, vol. 38, no. 8, pp. 1373–1379, 2003.

[12] R. Schoonhoven and D. F. Stegeman, "Models and analysis of compound nerve action potentials.," *Critical reviews in biomedical engineering*, vol. 19, pp. 47–111, Jan. 1991.

[13] G. Schalow, G. Zäch, and R. Warzok, "Classification of human peripheral nerve fibre groups by conduction velocity and nerve fibre diameter is preserved following spinal cord lesion," *Journal of the autonomic nervous system*, vol. 1838, no. 6, 1995.

[14] K. G. Pearson, R. B. Stein, and S. K. Malhotra, "Properties of action potentials from insect motor nerve fibres.," *The Journal of experimental biology*, vol. 53, pp. 299–316, Oct. 1970.

[15] M. Rahal, J. Winter, J. Taylor, and N. Donaldson, "An improved configuration for the reduction of EMG in electrode cuff recordings: a theoretical approach.," *IEEE transactions on bio-medical engineering*, vol. 47, pp. 1281–4, Sept. 2000.

[16] J. Taylor, N. Donaldson, and J. Winter, "Multiple-electrode nerve cuffs for low-velocity and velocity-selective neural recording," *Medical & Biological Engineering*, vol. 42, pp. 634–643, Sept. 2004.

[17] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "The theory of velocity selective neural recording: a study based on simulation.," *Medical & biological engineering & computing*, vol. 50, pp. 309–18, Mar. 2012.

[18] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "A summary of the theory of velocity selective neural recording.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2011, pp. 4649–52, Jan. 2011.

[19] A. Al-Shueli, C. Clarke, N. Donaldson, and J. Taylor, "Improved Signal Processing Methods for Velocity Selective Neural Recording Using Multi-Electrode Cuffs," *IEEE transactions on biomedical circuits and systems*, pp. 1–11, Sept. 2013.

[20] A. Back, "FIR and IIR synapses, a new neural network architecture for time series modeling," *Neural Computation*, 1991.

[21] E. A. Wan, "Time Series Prediction by Using a Connectionist Network with Internal Delay Lines," *Time Series Prediction*, pp. 195–217, 1994.

[22] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. Lang, "Phoneme recognition using time-delay neural networks," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 328–339, Mar. 1989.

[23] F. Karimi and S. R. Seydnejad, "Velocity Selective Neural Signal Recording Using a Space-Time Electrode Array.," *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, Dec. 2014.

[24] M. Abdallah, O. Rossel, S. Bernard, F. Soulier, and G. Cathebras, "Exploring the limit of ENG spatio-temporal filtering for velocity-selectivity," in *2015 7th Interna-*

tional IEEE/EMBS Conference on Neural Engineering (NER), pp. 585–588, IEEE, Apr. 2015.

[25] S. Gibson, J. W. Judy, and D. Markovic, "Spike Sorting: The First Step in Decoding the Brain," *IEEE Signal Processing Magazine*, vol. 29, pp. 124–143, Jan. 2012.

[26] K. Sameshima and L. A. Baccala, "Trends in multichannel neural ensemble recording instrumentation," in *Methods for Neural Ensemble Recordings* (M. A. L. Nicholelis, ed.), ch. 3, pp. 47–60, CRC, 1999.

[27] K. H. Kim and S. J. Kim, "Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier.," *IEEE transactions on bio-medical engineering*, vol. 47, pp. 1406–11, Oct. 2000.

[28] C. M. Stewart, S. D. Newlands, and A. A. Perachio, "Spike detection, characterization, and discrimination using feature analysis software written in LabVIEW.," *Computer methods and programs in biomedicine*, vol. 76, pp. 239–51, Dec. 2004.

[29] G. Zouridakis and D. Tam, "Identification of reliable spike templates in multi-unit extracellular recordings using fuzzy clustering," *Computer methods and programs in biomedicine*, vol. 61, pp. 91–98, 2000.

[30] S. Mukhopadhyay and G. C. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection.," *IEEE transactions on bio-medical engineering*, vol. 45, pp. 180–7, Feb. 1998.

[31] J. Kaiser, "On a simple algorithm to calculate the energy of a signal," *Acoustics, Speech, and Signal Processing*, vol. 2, no. 10, pp. 381–384, 1990.

[32] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials.," *Network (Bristol, England)*, vol. 9, pp. R53–78, Nov. 1998.

[33] M. Abeles and J. Goldstein, M.H., "Multispike train analysis," *Proceedings of the IEEE*, vol. 65, pp. 762–773, May 1977.

[34] J. Hartigan and P. Hartigan, "The dip test of unimodality," *The Annals of Statistics*, 1985.

[35] D. A. Adamos, E. K. Kosmidis, and G. Theophilidis, "Performance evaluation of PCA-based spike sorting algorithms.," *Computer methods and programs in biomedicine*, vol. 91, pp. 232–44, Sept. 2008.

[36] C. Zhang, X. Zhang, M. Q. Zhang, and Y. Li, "Neighbor number, valley seeking and clustering," *Pattern Recognition Letters*, vol. 28, pp. 173–180, Jan. 2007.

[37] R. Harrison, "A low-power integrated circuit for adaptive detection of action potentials in noisy signals," *Proceedings of the 25th Annual International Conference of*

the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), pp. 3325–3328, 2003.

Chapter 4

The Centroid Filter

4.1 The Need for Pulse-Time Information

Biological noise - as well as noise introduced by the recording electronics - produces a significant challenge in extracellular recordings. When noise is present, similar *action potentials* (APs) originating from different neurons may appear to share the same morphology, or APs from the same neuron may appear slightly different. One of the critical constraints of the spike sorting process is the alignment of individual APs before feature extraction is applied. Any temporal misalignment of APs at this time can have a detrimental effect on spike classification. The alignment process is often performed at higher sampling rates in order to minimise sampling jitter and the most common method of alignment is to align each AP to the point of *maximum slope* [1]. Other methods consider the point of maximum amplitude or the point (or points) at which the amplitude has been reduced by 3 dB. These methods typically employ only one or two samples from the entire AP and so are highly sensitive to broad-band noise that is typical of biological recordings [2]. Alignment to a metric that is derived from the *whole* AP rather than from a single point will be less susceptible to the effects of background noise in general, and it is this approach that is considered further.

A novel real time approach for spike alignment has been developed based on a *centre of gravity* filter. A new method called the *centroid filter* has been developed that provides an alternative to traditional spike alignment methods and substantially improves the resilience of the process to the effects of noise and sampling jitter. Validation of the new method is achieved by simulation using deterministic models of nerve signals (i.e. APs with added noise; where both correlated and uncorrelated noise models were considered) as well as results measured from an optimised hardware implementation. Comparisons have been made between these methods. A power efficient realisation of the centroid filter is described that operates in real time on a single Altera Max V *complex programmable logic device* (CPLD). The use of CPLDs for signal processing tasks is well documented and they are particularly well suited to the rapid analysis and assessment of novel techniques. A CPLD

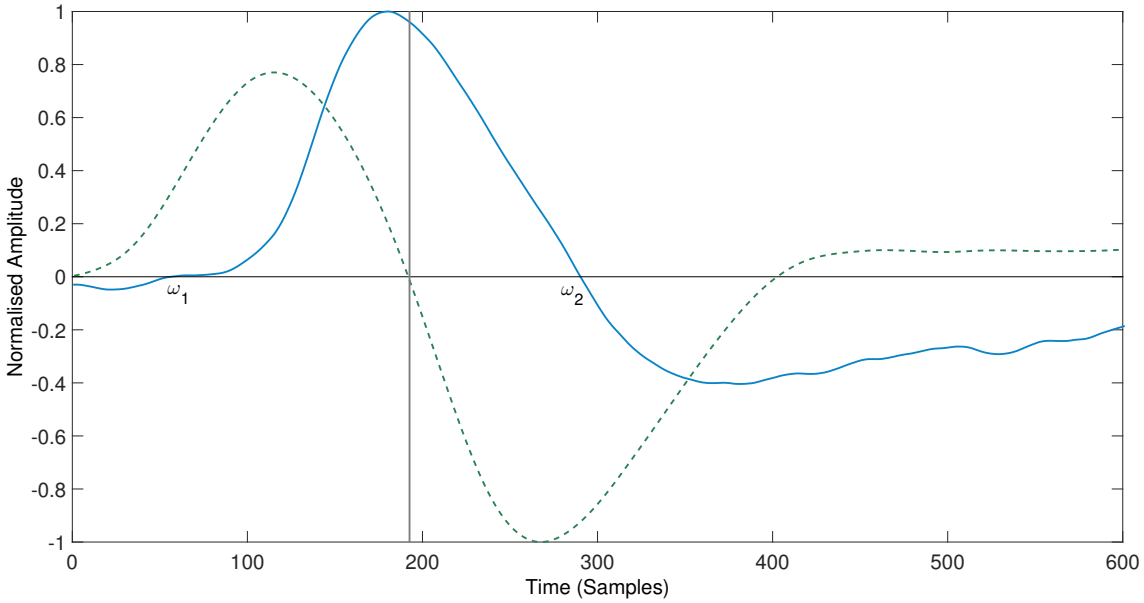


Figure 4.2.1: The application of the centroid filter to a single channel of data containing a realistic AP. In this case the width of the centroid filter was chosen to be $N = 100$ samples, or approximately the width of the positive phase of the AP. The solid line represents a discrete time version of the AP, and the dashed line the filter output $y[n]$. The vertical marker is set at the negative-going zero crossing of the filter output and is located at the centroid of the SFAP function.

is used in this work to evaluate the power and resource requirements for the proposed designs that are key metrics for implantable technologies [3].

4.2 The Centroid Filter

4.2.1 Operation

Within the spike sorting process the alignment of APs is crucial, and conventional single point methods can be strongly influenced by noise, introducing timing jitter and classification errors. A more robust and effective method that aligns waveforms based on an average of all of the samples within the AP can be derived; if the AP is considered to be a topological shape (i.e. a closed non self-intersecting polygon) then the central point of the AP in both time and amplitude can be considered to be the *centroid* [4]. Consider the AP shown in the solid trace of Figure 4.2.1. This was generated using the function in Equation 4.2.1, which is a first-order analytical model of a *single fibre action potential* (SFAP) that also includes a dependence on the fibre diameter; white Gaussian noise was added to this model [5, 6].

$$f(x) = Asin\left(\frac{x}{\tau_1}\right) e^{\frac{-x}{\tau_2}} \tag{4.2.1}$$

Where A is an amplitude constant that is a function of diameter and τ_1 & τ_2 control the duration of the AP, note that this function is only valid for $x > 0$.

The AP is bounded by the x axis and could be treated as a plane polygon with unit density per unit area. The centroid of the pulse calculated along the x axis can therefore be found by taking moments about the axis and dividing by the total area [7]:

$$C_x \equiv \frac{1}{\Upsilon} \int_{\omega_1}^{\omega_2} x f(x) dx \tag{4.2.2}$$

Where the centroid calculated along the x axis is C_x, Υ is the area under the AP, $(\omega_2 - \omega_1)$ is the width of the AP and $f(x)$ represents the AP function itself. Performing this integral on an unknown function numerically would be costly, typical approaches may involve fitting an interpolating function to the AP, where the integral of the interpolating function is well-known. Instead consider the convolution integral of $f(x)$ with another function $h(x)$, as represented in Equation 4.2.3.

$$y(x) = (h * f)(x) = \int_{-\infty}^{\infty} h(x - \lambda) f(\lambda) d\lambda \tag{4.2.3}$$

Note that if this integral is evaluated at the origin ($x = 0$), after a change of variables and a suitable choice of limits, it reduces to Equation 4.2.4:

$$y(0) = \int_{-\infty}^{\infty} h(-x) f(x) dx \tag{4.2.4}$$

Furthermore, if $h(x)$ is chosen to be kx, where k is a negative constant, then Equation 4.2.4 reduces to Equation 4.2.2. The importance of this lies in the fact that convolution can provide a very economical method to compute Equation 4.2.4, avoiding the use of the multipliers required for a direct numerical evaluation of the integral in Equation 4.2.2. To illustrate this process consider the test pulse $f(x)$ shown in Figure 4.2.2 (dashed top hat function of unit area). Evaluating the integral in Equation 4.2.2 where the limits of the integration are the dimensions of $f(x)$ along the x axis ($\omega_1 = -0.5$, $\omega_2 = 0.5$) the centroid can be shown to be at the origin ($C_x = 0$).

For the *convolution* method, as noted above, we choose $h(x)$ to be a linear function of x with a negative gradient passing through the origin. The width of the function h (i.e. the points at which $h = \pm 1$) is chosen to be greater than the width of f, for reasons that will become apparent shortly. In this example the constant N was chosen to be 1.5, as indicated in Figure 4.2.2 (Equation 4.2.5).

$$h(x) = \frac{-2x}{N} \tag{4.2.5}$$

For values of x where the two functions do not overlap the product of $(h * f)(x)$ is zero. Where overlap occurs, since both functions are discontinuous, a piecewise approach is taken. There are three phases to be considered, depending on the extent of the overlap of the two functions, as $h(x - \lambda)$ approaches $f(x)$ from the left. The first phase occurs as h

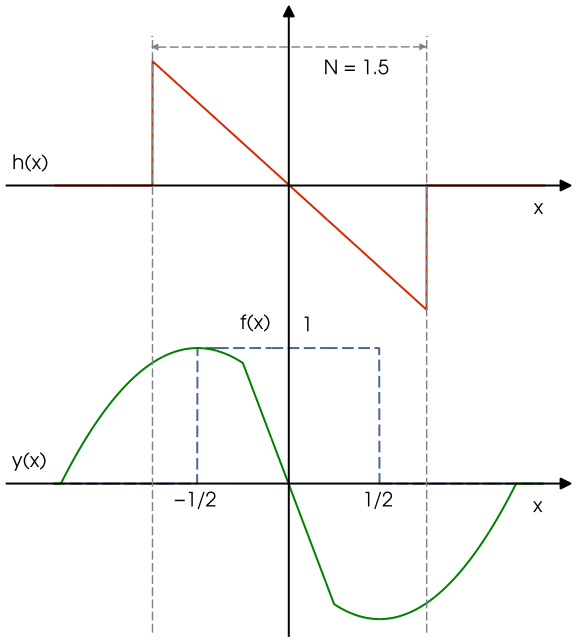


Figure 4.2.2: An example to illustrate the calculation of the centroid of an AP using the FIR filter $h(x)$. For clarity, h is represented by the continuous-time function shown in the upper plot in the figure while the 'top hat' function shown in the lower plot represents the AP. Application of the convolution and shift functions results in the output function $y(x)$ whose zero crossing corresponds with the centroid of the top hat function. Note $y(x)$ is shown in normalised form here for simplicity.

approaches ω_1, which is -0.5 in this case, and terminates when the leading edge of h is coincident with $\omega_2(0.5)$:

$$(h * f)^1(x) = \frac{-2}{N} \int_{\omega_1}^{x+0.5N} (x - \lambda) d\lambda = \frac{0.25N^2 - (x - \omega_1)^2}{N} \tag{4.2.6}$$

This function is a parabola symmetrical about the vertical axis $x = \omega_1$ (-0.5 in this case) and displaced vertically by $0.25N$, as indicated in normalised form in the lower plot in Figure 4.2.2. Similarly the third phase of the process ends when the trailing edge of h makes its last contact with ω_2 (i.e. 0.5) and is given by the following integral:

$$(h * f)^3(x) = \frac{-2}{N} \int_{x-0.5N}^{\omega_2} (x - \lambda) d\lambda = \frac{-0.25N^2 - (-x + \omega_2)^2}{N} \tag{4.2.7}$$

During phase 2, h completely encloses f since N was chosen to be greater than $|\omega_2 - \omega_1|$. In this phase the convolution integral is a linear function:

$$(h * f)^2(x) = \frac{-2}{N} \int_{\omega_1}^{\omega_2} (x - \lambda) d\lambda = \frac{(\omega_2 - \omega_1)}{N} [-2x + (\omega_2 + \omega_1)] \tag{4.2.8}$$

The function passes through the point $x = (\omega_2 + \omega_1)/2$, which is the centroid of $f(x)$. So in the example shown in Figure 4.2.2 , where $f(x)$ is symmetric about the origin in x ($\omega_1 = -\omega_2$), the convolution integral is simply a straight line passing through the origin with gradient $-2/N$. In addition, from Equations 4.2.6 to 4.2.8 it is easy to show that the

parabolas of phases 1 and 3 intersect with the linear phase 2 at the following points on the x axis:

$x_{12} = -0.5N - \omega_1$ (intersection of phases 1 & 2)
$x_{23} = 0.5N - \omega_2$ (intersection of phases 2 & 3)

For the example in Figure 4.2.2 these values of x are -0.25 and 0.25 respectively, as can be seen clearly in the figure. These equations also show the significance of the earlier observation that the function h must be wider than f. In the limit where the widths of the functions are equal, the intersection points x_{12} and x_{23} are both zero and the linear phase 2 disappears. In this case it is not possible to use the method to determine the centroid of $f(x)$. Since $h(x)$ can be realised as an *finite impulse response* (FIR) filter of length N, given that the relative length condition is satisfied, a single filtering operation on $f(x)$, plus zero crossing detection, is sufficient to calculate its centroid. Returning to the example shown in Figure 4.2.1, the result of the filter operation on a single realistic AP (solid line) is shown. The output from the filter crosses the origin at a point coincident with the centroid of the AP (dashed line).

The general form of the output expression $y[n]$ for an FIR filter of length N in the time domain with inputs $x[n]$ and coefficients b_i is as follows:

$$y[n] = b_0x[n] + b_1x[n-1] + ... + b_Nx[n-N] = \sum_{i=0}^{N-1} b_i.x[n-1] \tag{4.2.9}$$

To realise Equation 4.2.5 as an FIR filter we set:

$$b_i = \frac{-2i}{N} + 1 = mi + 1 \tag{4.2.10}$$

i.e. $b_0 = 1$ and $b_N = -1$ and $m = -2/N$ and where the constant (unity) term is required to make the filter realisable. However, substitution of Equation 4.2.10 into Equations 4.2.6 to 4.2.8 shows that the modified form of h introduces an output delay proportional to $N/2$ that offsets the position of the zero crossing. This can be compensated for by using a shift register of the same length to delay the input signal before processing.

4.2.2 Phase Analysis

The phase response of the centroid filter can be examined in more detail using standard filter analysis techniques. Considering a generalized-linear-phase system with a frequency response $H(e^{j\omega})$ and phase response:

$$\theta(\omega) = \beta - \alpha\omega \tag{4.2.11}$$

Where α and β are constants. By definition, the phase delay (the delay applied to each frequency) and the group delay (the average delay experienced by the composite signal at all frequencies) are given by:

$$T_{phase} = \frac{-\theta(\omega)}{\omega} = \frac{\beta}{\omega} - \alpha \tag{4.2.12}$$

$$T_{group} = \frac{-d\theta(\omega)}{d\omega} = \alpha \tag{4.2.13}$$

The impulse response of the centroid filter is anti-symmetric about either an integer or half an integer (depending if M is even or odd, also known as linear phase Type III/Type IV) as:

$$h(n) = \frac{-2n}{M} + 1 \tag{4.2.14}$$

Where:

$$h(n) = -h(M - n) \tag{4.2.15}$$

Where:

$$0 \leq n \leq N - 1 \tag{4.2.16}$$

Defining $M = N - 1$ and noting that at $n = M/2$:

$$h\left(\frac{M}{2}\right) = -h\left(M - \left(\frac{M}{2}\right)\right) = -h\left(\frac{M}{2}\right) \tag{4.2.17}$$

Thus:

$$h\left(\frac{M}{2}\right) = 0 \tag{4.2.18}$$

Proof of linear phase can be found by checking the Fourier transform of this type of filter:

$$H(e^{j\omega}) = \sum_{n=0}^{N} h(n)e^{j\omega n} \tag{4.2.19}$$

$$= e^{\frac{-j\omega M}{2}} \left[(-j) \sum_{n=0}^{\frac{M}{2}-1} h(n) 2sin\left(n - \frac{M}{2}\right)\omega \right] \tag{4.2.20}$$

The first term $e^{\frac{-j\omega M}{2}}$ gives a phase of $\frac{-\omega M}{2}$ to $H(e^{j\omega})$. Since $h(n)$ is real, the second term in the product above contributes a phase of $\frac{3\pi}{2}$ to $H(e^{j\omega})$. So the overall phase of $H(e^{j\omega})$ is:

$$\frac{-\omega M}{2} + \frac{3\pi}{2} \tag{4.2.21}$$

Using the definition of linear phase $\beta - \alpha\omega$ gives:

$$a = \frac{M}{2} \tag{4.2.22}$$

$$\beta = \frac{\pi}{2} \tag{4.2.23}$$

Returning to the definition of the group delay:

$$T_{group} = \alpha = \frac{M}{2} = \frac{N-1}{2} \tag{4.2.24}$$

Or, the number of samples delay applied (on average) to a composite signal over all frequencies is half the length of the filter, this result correlates to that found by substitution of Equation 4.2.10 into Equations 4.2.6 to 4.2.8.

4.2.3 Implementation as a Digital Filter

In a practical neural recording system it is preferable to perform as much signal processing as possible *in vivo*, minimising the need for invasive trans-cutaneous communications. If structures such as the centroid filter are to be used as part of a spike sorting based neural signal processing then attention must be given to the power consumption of the filter. It is possible to make a number of optimisations to the basic filter structure that greatly reduce the power required when implemented in a *very large scale integrated circuit* (VLSI) architecture such as an *field programmable gate array* (FPGA) or *application specific integrated circuit* (ASIC). In a direct form implementation, an FIR filter of length N samples requires N multipliers and N adders, and generally speaking multiplication will dominate the power budget. In this section a method is described that exploits the linear form of Equation 4.2.10 in order to avoid the use of multipliers as far as possible. Expanding Equation 4.2.9 and hence calculating the next output sample:

$$y[n] = b_0.x[n] + b_1.x[n+1] + ... + b_N.x[n-N] \tag{4.2.25}$$

$$y[n+1] = b_0.x[n+1] + ... + b_N.x[n-N+1] = \sum_{i=0}^{N} b_i.x[n+1-i] \tag{4.2.26}$$

Recalling that $b_0 = 1$ and $b_N = -1$, then, from Equation 4.2.10:

$$b_i = mi + 1 \tag{4.2.27}$$

The following recurrence relationship between the coefficients can be derived:

$$b_i - b_{i-1} = m \tag{4.2.28}$$

Using this expression, Equation 4.2.25 for the next output sample can be rearranged as follows:

$$y[n+1] = y[n] + \left\{ m \sum_{i=0}^{N-1} x[n-i] \right\} + x[n+1] - x[n-N] \tag{4.2.29}$$

The computation of $y[n+1]$ in this way requires $N+3$ additions and one multiplication. Furthermore, if m is a power of 2 the multiplication can be replaced by a left shift. Also, the computation of the summation term in Equation 4.2.29 can be accomplished by using a rolling sum expressed as follows where $\omega[n]$ is the rolling sum:

$$\omega[n] = m \sum_{i=0}^{N-1} x[n-i] \tag{4.2.30}$$

Equation 4.2.30 may be expanded as follows:

$$\omega[n+1] = \omega[n] + m\{x[n+1] - x[n-N-1]\} \tag{4.2.31}$$

Substitution of Equation 4.2.30 and Equation 4.2.31 back into Equation 4.2.29 results in the following optimised expression for $y[n+1]$:

$$y[n+1] = y[n] + \omega[n] + m\{x[n+1] - x[n-N-1]\} + x[n+1] - x[n-N] \tag{4.2.32}$$

The computational effort has been reduced drastically compared to a direct implementation of an FIR filter and an implementation of Equation 4.2.32 requires *one* multiplier and *five* adders regardless of the length of the filter (the earlier remark about m being chosen as a power of 2 still applies). The shift register used to store the previous samples is the only component that is dependent on the order of the filter.

4.3 Results

4.3.1 Power & Area Measurements

An analysis of the power and resource requirements for the proposed centroid filter was performed using dedicated hardware implementations constructed using (a) an FPGA (Altera Cyclone II EP2C35F672C6N) and (b) a CPLD (Altera MAX V 5M570ZF256C5N). These devices were chosen as rapid prototyping platforms both because of their low cost compared to a custom ASIC and because at lower sampling rates an FPGA or CPLD can consume less power than an equivalent processor-based implementation [8]. Furthermore the relatively low sampling rates associated with ENG, typically less than 100 kS/s, may allow a reduction in core operating voltages and thus a saving in static power consumption in an FPGA [9]. Additionally the use of a MAX V CPLD allowed for separate external regulation of the core and I/O voltage busses. This configuration permits more accurate power consumption measurements as the two power domains can be easily separated. The designs were produced using the SystemVerilog hardware description language and *register transfer level* (RTL) verification was performed using the QuestaSim environment (*QuestaSim 10.2c, Mentor Graphics Inc, Oregon, USA*).

The synthesis tools (*Quartus 13.0, Altera Corporation, Calif. USA*) reported an initial maximum clock rate for the centroid filter of 25 MHz. The designed clock rate was chosen to be 500 kHz, representative of the fastest sampling rate used in recent acute experiments [10]. Power measurements were made using a current sense resistor (1Ω) connected upstream and in-line with the DC power supply. Baseline power measurements were made using clock gating to isolate the filter structure, as well as to estimate the power consumed by the I/O and clock propagation circuitry. The CPLD implementation utilized separate core and I/O power supplies, both at 3.3 V, for more accurate measurements. The power requirements were measured at a clock frequency of 500 kHz. Resource measurements were taken from the synthesis tools after optimisation and fitting of the design. An overview of the resource and power requirements achieved for both methods, is provided in Table 4.3.1. Power analysis tools (*Powerplay Power Analyser, Altera Corporation*) were used to predict the power consumption of the designs and the predicted values were in close agreement with measured data.

Table 4.3.1: Comparative device utilization summary and power consumption for the centroid filter in an FPGA and a CPLD $f_{clk} = 500$ kHz.

Device	Parameter	Result
CPLD - 5M570ZF256C5N	Total Equivalent Gate Count	442
	Current Consumption	0.09 mA
	Total Power Consumption (Measured)	0.17 mW
	Total Power Consumption (Estimated)	0.15 mW
FPGA - EP2C35F672C6N	Total Equivalent Gate Count	442
	Current Consumption	34.5 mA
	Total Power Consumption (Measured)	62.2 mW
	Total Power Consumption (Estimated)	61.5 mW

4.3.2 Verification

Test pattern waveforms were generated in MATLAB using the SFAP model of Equation 4.2.1 before being sampled and transformed into a $Qs0{:}7$ fixed point number format. The RTL codes used to produce designs for the two devices were identical. In addition to the sample memories a verification block was included within the design. This block contained the expected output from both filters and performed simple online comparisons on a sample-by-sample basis. The data files for both the sample memory and the verification block were transferred using a custom *joint test action group* (JTAG) programming interface into the on-chip synchronous memories - implemented in 4K *random access memory* (RAM) blocks - of both the FPGA and the CPLD. A memory pointer was driven from a variable down sampled clock (system clock speed was 500 kHz, down sampled from 10 MHz), this pointer was used to address the circular memory buffer, the returned samples being fed into the centroid filter. The clock control circuitry was used to activate individual modules to obtain accurate power measurements. The verification circuity confirmed correct operation of the centroid filter up to a maximum clock frequency of 24.5 MHz at which point the expected output deviated from the observed output, this frequency is in agreement with the predicted maximum clock frequency of 25 MHz.

4.3.3 Comparison to Traditional Methods in Suboptimal Noise Environments

In order to evaluate the performance of the centroid filter when applied to AP alignment a number of simulations were performed in which APs were aligned using *four* metrics: the centroid, the maximum value, the point of maximal slope and the -3 dB points. Each point was first calculated in a noise free environment to obtain the reference value. Different types of noise where then applied and the study examined the statistical effect of noise on each of the metrics.

Table 4.3.2: Model parameters that define the SFAP functions computed using Equation 4.3.1.

Axon Diameter (μm)	**A**	τ_1(ms)	τ_2(ms)
5	2.42	0.175	0.25
7	2.65	0.120	0.15
9	2.73	0.093	0.11
11	2.73	0.080	0.096
13	2.79	0.078	0.092
15	2.80	0.076	0.089
19	2.89	0.072	0.084

Single Fibre Action Potential Model

The SFAP model used for generating APs has already been described in Equation 4.2.1. It is based on a damped sinusoidal function and was proposed by [5, 11]. It is presented here with a change of variables from x to t representative of discrete time sampling:

$$f(t) = D^2 A sin\left(\frac{t}{\tau_1}\right) e^{\frac{-t}{\tau_2}} \tag{4.3.1}$$

Where D is the diameter of the axon and the parameters A, τ_1 and τ_2 determine the amplitude, rising edge rate and duration of each AP, respectively. The function is only valid for $t > 0$. The parameters used for a range of different axonal diameters are listed in Table 4.3.2, having been determined from experimental recordings made from the sciatic nerve of *frog* [11]. The parameters are arranged so that the SFAP amplitude and duration are both a function of the diameter of the axon, for the purposes of simulation a range of different diameters were considered in order to provide a broad selection of different SFAP waveforms. Fig. 4.3.1 shows the seven SFAPs generated using this model, where the parameters have been extracted from Table 4.3.2. The fibre diameters shown in Figure 4.3.1 are: $5, 7, 9, 11, 13, 15$ & 19 μm. The corresponding conduction velocities are: $9.3, 13, 16.7, 20.4, 24.2, 28$ & 35.3 m/s respectively.

Noise Models

In each simulation random noise was added to noiseless APs to produce a signal with a specific, controllable *signal-to-noise ratio* (SNR). The following standard formula was employed to calculate SNR:

$$SNR(db) = 10log\left(\frac{P_{signal}}{P_{noise}}\right) = 10log\left(\frac{mean^2_{signal} + \sigma^2_{signal}}{mean^2_{noise} + \sigma^2_{noise}}\right) \tag{4.3.2}$$

Where P_{signal} and P_{Noise} are the signal and noise powers and σ^2_{signal} and σ^2_{noise} are the variances respectively.

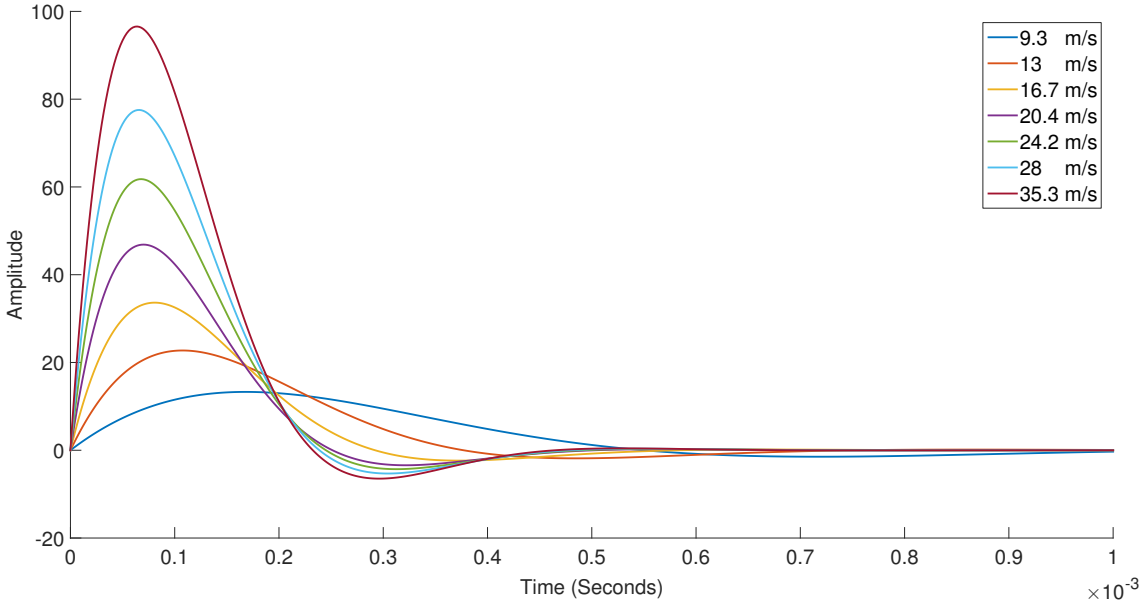


Figure 4.3.1: Seven different APs generated using the damped sinusoid model given in Equation 4.2.1. The diameters are 5 μm (blue), 7 μm (red), 9 μm (yellow), 11 μm (mauve), 13 μm (green), 15 μm (blue) & 19 μm (burgandy) corresponding to conduction velocities of 9.3 m/s, 13 m/s, 16.7 m/s, 20.4 m/s, 24.2 m/s, 28 m/s & 35.3 m/s respectively.

Both *additive white Gaussian noise* (AWGN) and *correlated stochastic noise* (CSN) processes were considered since both have relevance to the neural recording environment. For example AWGN is frequently used in theoretical studies of neural recordings because myelinated axons are considered to be electrically isolated and therefore have no synaptic interactions within a nerve [12]. In theory, therefore, background noise in recordings made from hook or suction electrodes of large myelinated nerves could be approximated by a Gaussian noise process. In practice, however, *ephaptic* interactions between axons and crosstalk between amplifiers and recording equipment introduce various levels of correlation [5]. Therefore a CSN noise model was also considered and used to describe the background activity. The noise process was correlated in time and it was generated using a dynamic *Ornstein-Uhlenbeck* (OU) process described by Equation 4.3.3 [13]:

$$OU_{t+dt} = OU_t - \frac{OU_t}{\tau}dt + d\Omega_t \tag{4.3.3}$$

Where τ is the time constant of the process and was set to be 0.01, dt is the simulation time step and Ω_t denotes a Wiener process.

Results with Correlated and Uncorrelated Noise

In the simulations described in this section both types of noise processes were applied to simulated APs *separately* using the following method. A reference AP was generated representative of that expected from a myelinated axon of diameter 15 μm at a temperature of 37 degrees Celsius. The reference AP was sampled at 500 kHz by multiplication with an

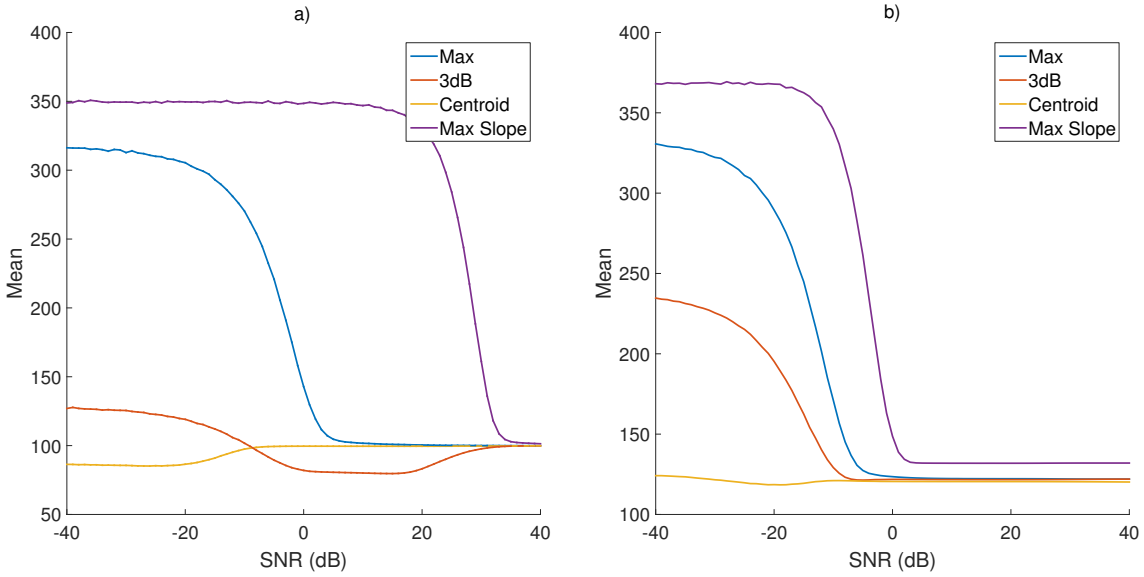


Figure 4.3.2: The mean locations for the pulse parameters in the presence of additive white Gaussian noise with SNR from 40 dB to −40 dB. **Left inset** - no filtering is applied to the noise. **Right inset** - a low-pass filter with a cut-off frequency of 10 kHz was used which is representative of the first stage of the spike sorting process. The simulated AP is representative of an axon with diameter 15 μm at a temperature of 37 degrees Celsius. The plots have been normalised to a theoretical mean of 100.

impulse train. Using this reference AP the ideal values of the four pulse parameters were measured and stored, (a) the point of maximum slope, (b) the maximum value, (c) the midpoint of the −3 dB points and (d) the centroid. Independently noise was then added to the reference waveform representative of SNR values in the range +40 dB to −40 dB in steps of −1 dB. For each value of SNR the experiment was repeated 100,000 times and the pulse parameters (a to d) were computed each time and stored for statistical analysis. Both the mean and the standard deviation of the pulse parameters were computed for each value of SNR with both types of noise, the values were normalised to a common reference for the purposes of comparison and presentation.

The normalized plots shown in Figures 4.3.2 and 4.3.3 were obtained by applying these methods using AWGN as the noise source. The mean locations and standard deviations of the pulse parameters are plotted using the four measures listed above, which will be referred to as (a), (b), (c) and (d) respectively. For the case of AWGN, in addition to applying the noise directly, the signal was low-pass filtered representative of the first part of the spike sorting process. An 8^{th} order Butterworth digital *infinite impulse response* (IIR) filter with a cut-off frequency of 10 kHz was employed. Figures 4.3.2 and 4.3.3 are split into two pairs (4.3.2a, 4.3.2b, 4.3.3a & 4.3.3b) to accommodate this comparison.

The main feature of these results (i.e. for both mean and standard deviation) is that pulse location based on maximum slope (a) appears to be the most susceptible to noise, degrading severely with an SNR less than about 30 dB. This is followed by the method based on the maximum value of the pulse (b), which fails with an SNR of approximately

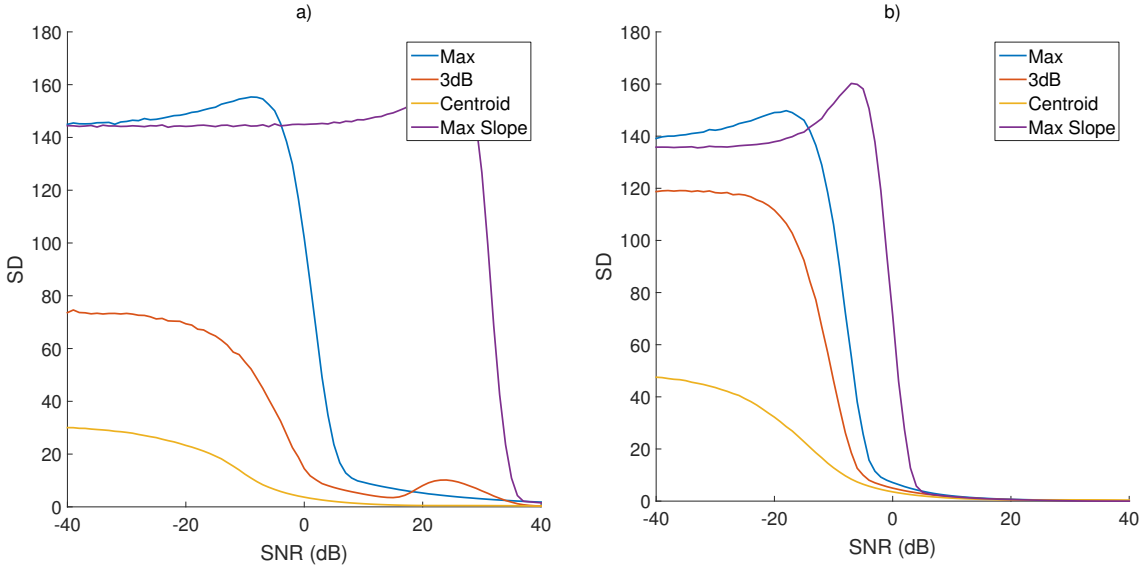


Figure 4.3.3: The standard deviations of the pulse parameters in the presence of additive white Gaussian noise with SNR from 40 dB to −40 dB. **Left inset** - no filtering is applied to the noise. **Right inset** - a low-pass filter with a cut-off frequency of 10 kHz was used which is representative of the first stage of the spike sorting process. The simulated AP is representative of an axon with diameter 15 μm at a temperature of 37 degrees Celsius.

0 dB. The methods based on the mean of the −3 dB points (c) and the centroid (d) perform much better, operating well with SNRs of about −10 dB. In these simulations the centroid method always outperformed the others.

In the second case, white (i.e. uncorrelated) noise was replaced by correlated noise generated using the OU model defined above. The experiments were repeated without additional filtering and the results are shown in Figures 4.3.4 and 4.3.5. The main effect of using correlated noise was that methods (a) and (b), which performed worst under AWGN (especially the case where filtering was not employed), performed much better using OU noise. The overall effect was that the spread in performance of all four methods was much less than when AWGN was used, for both the mean and SD measures. This suggests that single point temporal measures such as methods (a) and (b) are particularly sensitive to high frequency noise components such as are present in AWGN. In addition both these methods rely on determining *turning points* in the time record, emphasising their vulnerability to noise in general and high frequency components in particular.

4.4 Selectivity Improvements in Velocity Selective Recording

A further application of the centroid filter can be found within *velocity selective recording* (VSR) - the improvement in velocity selectivity via the artificial reduction in AP duration. As discussed in Chapter 3 one of the more successful methods for artificially increasing the velocity selectivity of a VSR system is the inclusion of a band-pass filter after the delay-

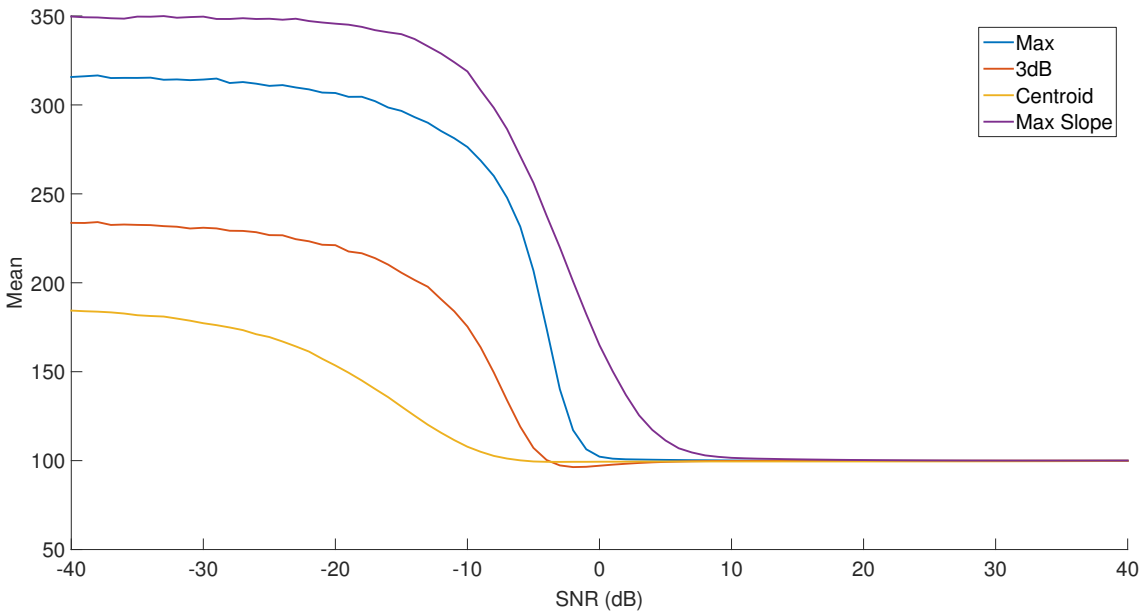


Figure 4.3.4: The mean locations for the pulse parameters in the presence of Orstein-Uhlenbeck noise with SNR from 40 dB to −40 dB. The simulated AP is representative of an axon with diameter 15 μm at a temperature of 37 degrees Celsius. The plots have been normalised to a theoretical mean of 100.

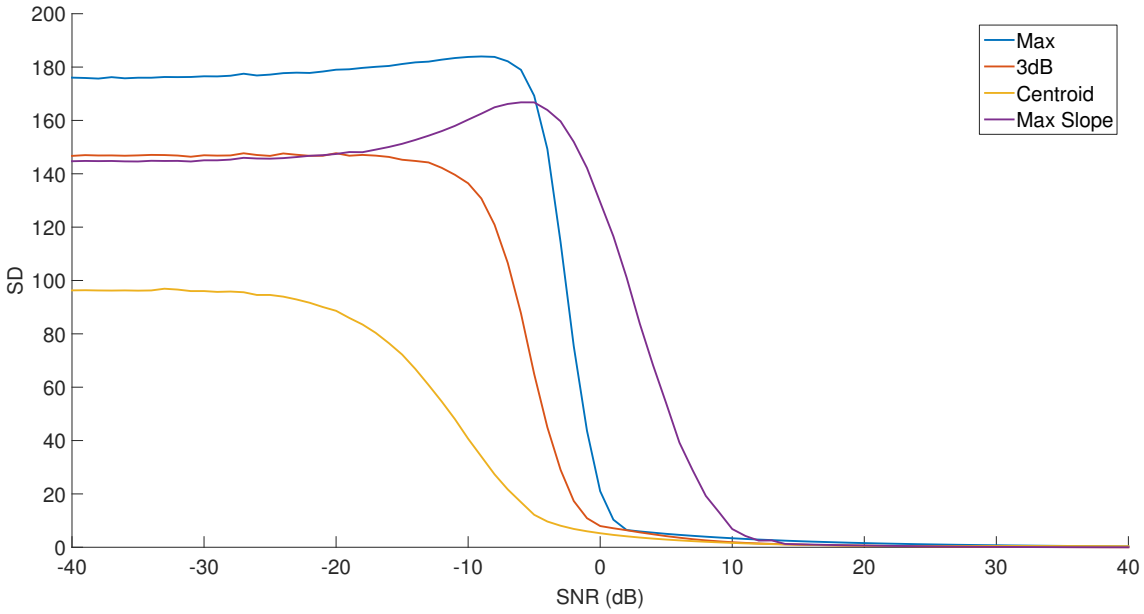


Figure 4.3.5: The standard deviations of the pulse parameters in the presence of Orstein-Uhlenbeck noise with SNR from 40 dB to −40 dB. The simulated SFAP is representative of an axon with diameter 15 μm at a temperature of 37 degrees Celsius.

and-add process. The basic premise of band-pass filtering is that the APs are replaced by bursts of damped sinewaves, with characteristics dependant on the filter structures. The delay-and-add process is now effectively matching these damped bursts, which have a centre lobe that is narrower in time than that of the original AP. One of the limitations of the band-pass filter method however is that the *side-lobes* create images within the velocity spectra, and as the centre frequency (and thus velocity selectivity) of the band-pass filter is increased the relative amplitude between the peak velocity and the side-lobe images within the spectra decreases. This issue is illustrated in Figure 4.4.3 (left inset) where images are observable for centre frequencies of 8 kHz and upwards.

Alternatively the AP duration can be reduced by the application of a modified centroid filter, a *width reduction envelope* (WRE) that can be easily derived from Equation 4.2.32. The WRE is computed as follows:

$$WRE[n+1] = 1 - |y[n+1]| \tag{4.4.1}$$

The WRE consists of two inverted sections of a parabola (corresponding to phases 1 and 3 in Figure 4.2.2) and a triangular notch (corresponding to phase 2 in Figure 4.2.2). The original input signal is multiplied by the triangular notch phase of the WRE so that each AP is narrowed in time by the envelope only at points that correspond to the centroid. This process is illustrated in the upper trace in Figure 4.4.1 for a filter of order $N = 2$ where the centroid is located at zero. Assuming the fundamental assumption that the width of the AP is less than N, multiplication of the two functions will result in a narrower AP. Clearly the exact degree of width reduction also depends on the form of the AP and so cannot be easily stated in general. However, for the example of the top hat test function, the width reduction can be readily determined, as illustrated in the lower traces of Figure 4.4.1, showing the original top hat function (dashed line) and the result of multiplying this function with the WRE (solid line). The width of the top hat function (or the AP) has been reduced by a factor of approximately 41.4% as observed at the -3 dB points. It can be shown that for values of N larger than the intrinsic width of the top hat the resulting width will be:

$$\omega_{new} = 2\left(\frac{2(V_p) - 1.414}{N}\right) \tag{4.4.2}$$

Where ω_{new} is the width of the top hat function as observed at the -3 dB points after the narrowing process and V_p is the amplitude of the original waveform. In a practical recording system the amplitude of the AP (V_p) will be normalised in order to maximise the overall dynamic range of the system, the parameter N can be selected so that it is approximately the width of the AP in order to maximise the pulse reshaping effect. Standard delay-and-add techniques can then be applied to the time domain data to compute the velocity spectrum using the new reshaped pulses.

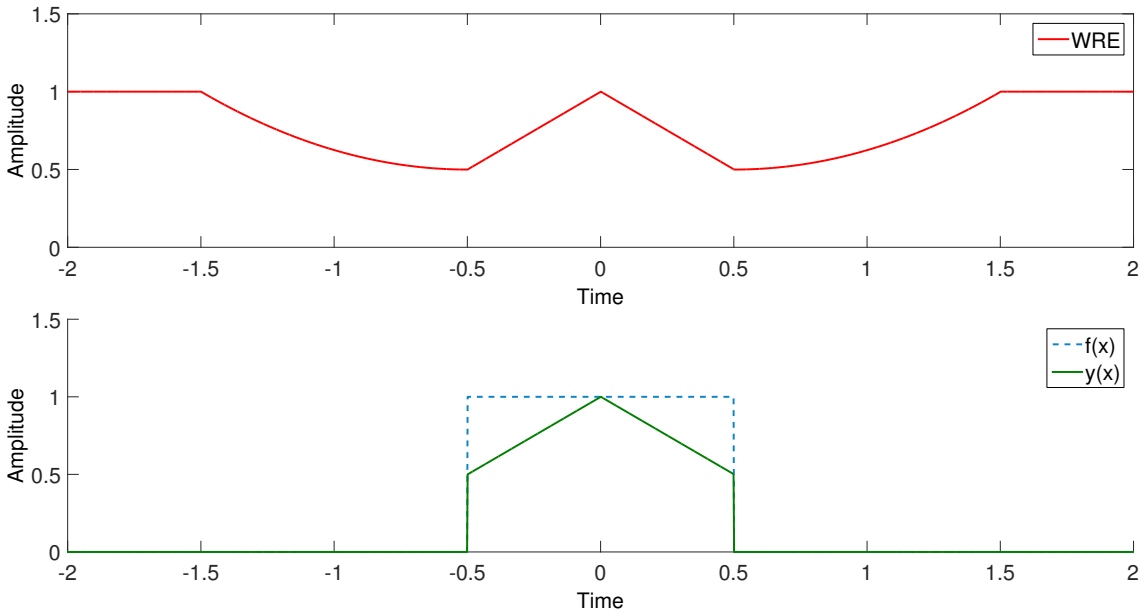


Figure 4.4.1: Top Inset - The WRE computed using Equation 4.4.1 for the top hat test function and a centroid filter of width $N = 2$. **Bottom Inset -** The original top hat test function (dashed line) and the reshaped (narrowed) top hat function after multiplication with the WRE (solid line). The top hat function has been narrowed by 41.4% of its original width as observed at the -3 dB points.

4.4.1 Simulated Selectivity Improvements

A simulation study was performed in order to compare the relative increase in velocity selectivity available using band-pass filters and the modified centroid filter. Ten channels of data containing a single SFAP were generated using the model given in Equation 4.2.2, which has been proposed as a suitable approximation for the simulation of mammalian ENG [14]. Artificial time domain recordings were generated using MATLAB for a SFAP with a conduction velocity of 30 m/s in a *multiple electrode cuff* (MEC) with total length 3 cm containing 11 equally spaced electrode contacts grouped to realise 9 tripolar channels of data. The sample rate was 100 kS/s and white Gaussian noise was added independently to each channel to provide a total channel SNR of 10 dB.

A single channel of simulated data is shown in Figure 4.4.2, the dashed curve is the original AP and the solid curve is the same AP after the application of a WRE computed from a centroid filter with order $N = 100$ samples, approximately the width of the AP. As shown in Figure 4.4.2 the width of the AP has been reduced but importantly the location in time (the centroid) of the AP has not changed. From this data an *intrinsic velocity spectrum* (IVS) was calculated using the delay-and-process (Figure 4.4.3, right hand inset) and shows a clear increase in velocity selectivity. For the simulated data set the intrinsic selectivity is 3.53 whereas the centroid filtered spectrum has a velocity selectivity of 8.57. This represents an increase of approximately 2.5 over the intrinsic value.

The MATLAB simulations were repeated with a single band-pass filter of centre frequency

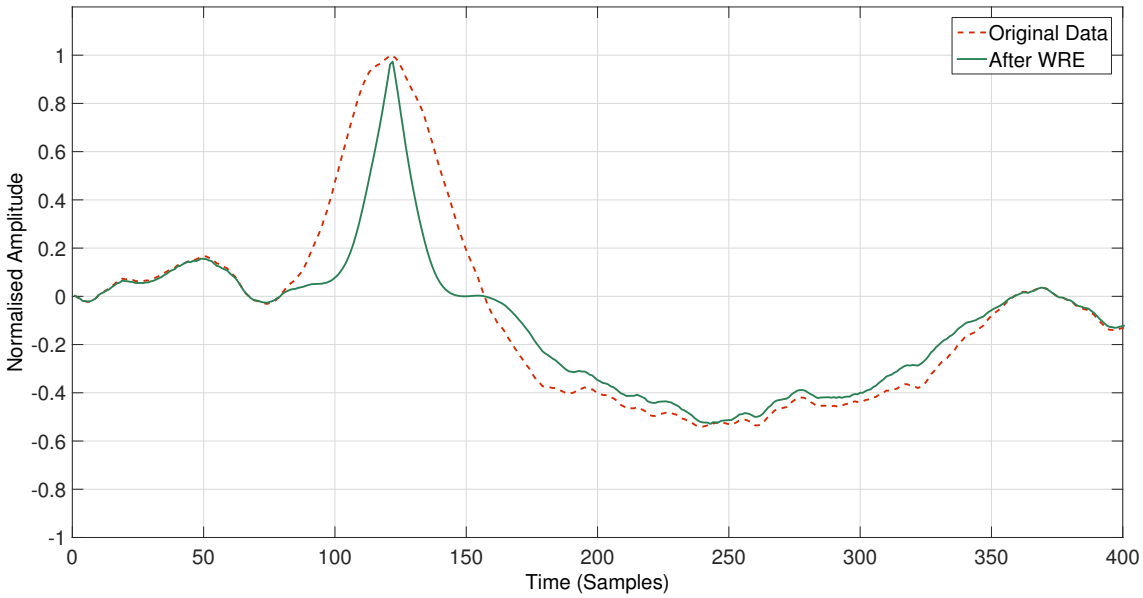


Figure 4.4.2: Time domain recording of a single AP with additive white Gaussian noise of SNR 10 dB (dashed line) and the same AP after width reduction has been applied using the WRE (solid line), order of the centroid filter $N = 100$ samples.

f_0 placed at the output of the delay-and-add process (see Figure 3.2.2). This is equivalent to placing a filter on the output of each channel due to the linearity of the system [15]. The band-pass filters were 8^{th} order Butterworth filters with centre frequencies of 1 kHz, 2 kHz, 4 kHz, 8 kHz, 16 kHz and 32 kHz respectively. The relative bandwidth of each filter was 20%. The velocity spectra for each filter (Figure 4.4.3 left hand inset) illustrate the increase in velocity selectivity as the centre frequency f_0 is increased. Q_v ranges from 0.80 to 33.3 and is approximately proportional to f_0. Note the spurious image responses that are visible in the three spectra with $f_0 = 8$ kHz and above. Note also that in practice the available bandwidth in biological recordings is typically less than about 10 kHz, and so the maximum value of Q_v obtainable using this method is about 7. Whereas the centroid filter increases Q_v to 8.57. On this basis, the centroid filter method outperforms band-pass filters, especially when the cost of hardware implementation is considered.

In order to examine the differences in implementation costs an 8^{th} order Butterworth band-pass filter was implemented in an Altera Cyclone II EP2C35F672C6N with a centre frequency of 8 kHz and a relative bandwidth of 20% in agreement with theoretical studies [15]. The centre frequency of the band-pass filter was chosen to be 8 kHz so that the velocity selectivity (6.7) was as similar as possible to the centroid filter (8.57). The band-pass filter was implemented using the Direct-Form II method with second order sections. The centroid filter had order $N = 100$ samples and was implemented using the method described in Section 4.2.3. Additionally the input samples were constrained to be signed fixed point 8-bit values and the output values were signed 32-bit. The designs were produced using the SystemVerilog hardware description language and RTL verification was performed using the QuestaSim environment (*QuestaSim 10.2c, Mentor Graphics Inc,*

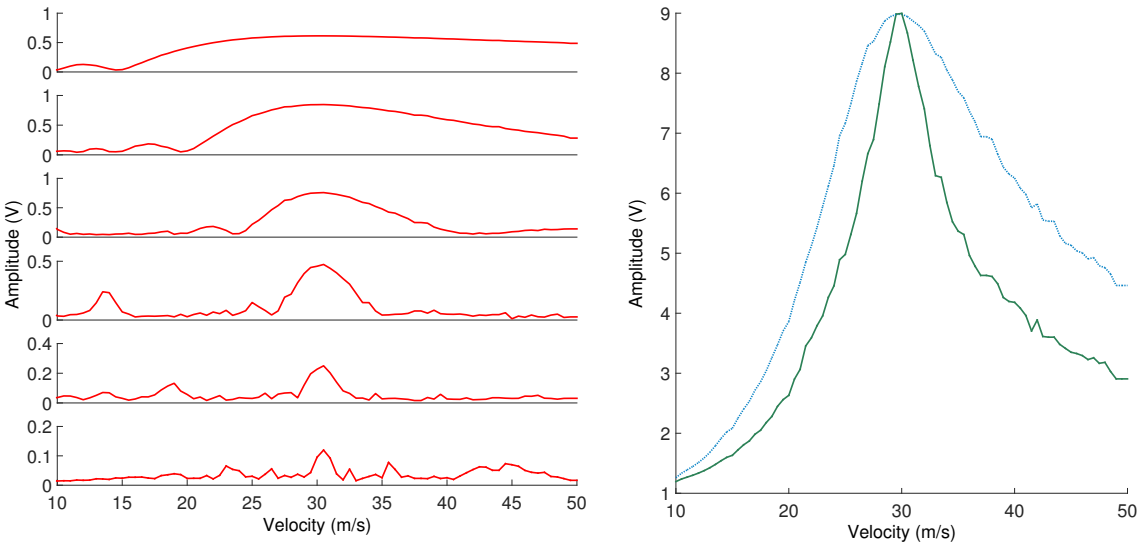


Figure 4.4.3: Left Inset - Band-pass filtered velocity spectra for a single AP propagating at 30 m/s. Filter centre frequencies and velocity selectivities are: a) 1 kHz - 0.8 b) 2 kHz - 2.03 c) 4 kHz - 4.06 d) 8 kHz - 6.7 e) 16 kHz - 15.2 f) 32 kHz - 33.2. **Right Inset -** The IVS (dashed line) with $Q_v = 3.53$ and the centroid filtered velocity spectrum with $Q_v = 8.57$ representative of a 2.5 times increase in selectivity.

Oregon, USA).

The synthesis tools (*Quartus 13.0, Altera Corporation, Calif. USA*) reported an initial maximum clock rate for the band-pass filter of 17.43 MHz and a maximum clock rate for the centroid filter of 25 MHz. The designed clock rate was chosen to be 500 kHz, representative of the fastest sampling rate used in recent experiments [10]. Power measurements were made using a current sense resistor (1Ω) connected upstream and in-line with the DC power supply. Baseline power measurements were made using clock gating to isolate the two filter structures, as well as to estimate the power consumed by the I/O and clock propagation circuitry. Resource measurements were taken from the synthesis tools after optimisation and fitting of the design. An overview of the resource and power requirements, as well as the velocity selectivity achieved for both methods, is provided in Table 4.4.1. Power analysis tools (*Powerplay Power Analyser, Altera Corporation*) were used to predict the power consumption of the designs and the predicted values were in close agreement with measured data. The band-pass filter consumed approximately 10 times more power than the centroid filter.

Measured results (see Table 4.4.1) show that the centroid filter uses approximately 90% less power than an equivalent band-pass filter, for a similar level of velocity selectivity. However, one of the benefits of the band-pass filter method is that only a single filter is required per delay-and-add module (due to the linearity of the system). In a typical VSR recording setup there may be up to 9 channels of raw data to process and up to this limit it would be power (and area) efficient to use a bank of centroid filters over the band-pass filter. One drawback of the band-pass filter method is the introduction of spurious images

Table 4.4.1: Comparative device utilization summary and power consumption f_{clk} = 500 kHz.

Method	Parameter	Result
Band-pass Filter	Total Equivalent Gate Count	2533
	Current Consumption	343.7 mA
	Total Power Consumption (Measured)	618 mW
	Total Power Consumption (Estimated)	601 mW
Centroid Filter	Total Equivalent Gate Count	442
	Current Consumption	34.5 mA
	Total Power Consumption (Measured)	62.2 mW
	Total Power Consumption (Estimated)	61.5 mW

within the velocity domain, Figure 4.4.3 (left inset) illustrates this problem. As the centre frequency and thus selectivity of the filter is increased the amplitudes of the side lobes within the time domain are increased, this produces the spurious peaks that are seen at lower velocities. Of course if a band-pass filter is fitted to a specific delay-and-add filter for a single velocity then these spurious images could be rejected as being out of band, however this has yet to be demonstrated in a practical system.

Placed in the context of the total recording system, a typical three electrode system manufactured in a 0.8 μm *complimentary metal oxide semiconductor* (CMOS) technology consumes approximately 105 mW of power for the low noise amplifiers and front end electronics and approximately 129 mW of power for data conversion and digitisation [16]. The addition of high precision band-pass filters that may consume up to 600 mW is therefore a significant addition to the total system power and centroid filters may be of significant benefit in a power critical environment such as a fully implanted neural recording system.

The work presented here has focussed on neural recordings where the APs are distinct and visible within the time domain. This is generally the case with, for example, acute *in vivo* recordings made using hook electrodes [10], but it often not the case with chronically implanted electrode structures. Distinct APs have been recording using microchannel electrode structures and it is possible to extend such a structure to include the multiple electrodes required to perform VSR filtering.

4.5 Discussion

4.5.1 General Effects of Noise on Spike Alignment

As outlined in Chapter 3, neural recording methods have developed greatly in recent years with the development of multi-electrode methodologies for a wide variety of applications. Of these methods, spike sorting is amongst the most popular but unfortunately also one

of the most strongly influenced by noise [17]. This is because the methods employed to align the position of individual spikes (APs) in time tend to depend on single point measurements of each spike and are therefore particularly sensitive to the effects of noise processes. This sensitivity was demonstrated in the simulations reported in this chapter for both uncorrelated and correlated additive noise sources. The basic approach adopted here is therefore to propose a method that does not rely on single point determinations of spike locations, but rather employing a measure that in some sense is an average of the whole AP. As has been demonstrated, this has the effect of reducing the sensitivity of the spike position measurement to additive noise, at least for uncorrelated sources.

4.5.2 Comparison of Correlated and Uncorrelated Noise

It is tempting to assume that because *myelinated* axons are generally considered to be electrically isolated and therefore to have no synaptic interactions within a nerve that all noise sources can be considered to be uncorrelated (AWGN) [12]. In practice, however, various interactions between axons and crosstalk between amplifiers and recording equipment introduce some correlation [5]. Therefore a correlated noise model was also considered in the simulation study, based on the OU process [13]. The effect of this change was dramatic on the two methods of spike position determination based on single points in time. These were (a) the point of maximum AP slope and (b) the actual maximum point. This is not a surprising result because in addition to being single point measures these methods require the calculation of a turning point in time, suggesting an enhanced sensitivity to high frequency noise components. This view was confirmed by low pass filtering the signal after the application of AWGN. This is standard practice in spike sorting and the result was comparable to the effect of changing to a correlated source, which is naturally band limited (Figures 4.3.4 & 4.3.5).

4.5.3 Benefits of the Centroid Filter

In all cases, and independent of the type of noise applied to the recordings, the centroid proved to be a significantly more reliable method for time aligning the APs. Intuitively the centroid of the AP is a function of the energy contained within the AP itself, the addition of AWGN to a first order approximation does not change the amount of energy in the AP and so will not affect the position of the centroid. Of course this simplification does not hold when considering finite time sequences or band limited noise such as the OU process considered in this study.

In its most basic form computing the centroid introduces significant overhead, especially in comparison to the single sample measures. However it has been shown that the computation can be performed using minimal effort using a novel FIR filter structure (considering the centroid filter to be an FIR filter it has order N, the duration of the AP). It has been shown that the application of a low pass filter to the raw signal improves the performance

of the single sample methods for AP alignment and in this study an 8^{th} order Butterworth IIR low pass filter with an upper cut-off frequency of 10 kHz was employed. The choice of low pass filter order is driven largely by the Nyquist rate, which in turn limits the filter order (high filter orders with low cut-off frequencies produce unstable filters). An 8^{th} order filter was found to provide good stability and well as sharp transitions from the pass band to the stop band. When using signal conditioning filters in this way careful attention must be paid to the phase response of the filter, in a noise free environment the APs will have broadly similar frequency spectra and so the effect of a filter with non-linear phase will be minimal. In practice the spectra of each AP is a function of many independent variables and will vary from one AP to another. A filter with non-linear phase will therefore have some effect on the overall morphology of each AP unless the filter is designed in such a way that the phase response was homogenous within the pass band region. It was shown in Figure 4.3.2 (b), for example, that a low pass filter with an upper cut-off frequency of 10 kHz significantly shifts the point of maximal slope even in the noise free case, from 100 samples to 150 samples in this case.

When considering the implementation costs then the centroid filter, using only a single multiplier, will significantly outperform a direct form IIR filter of order greater than unity. Of course many different implementation forms of IIR filters are available but in general the number of multipliers required scales linearly with the order of the filter. The centroid filter is a low power method for AP alignment that shows considerable resilience to noise even in SNRs less than -10 dB. This method may have applications in a wide range of areas wherever phase information between different pulses is considered, for example in phase sensitive communication or radar systems.

4.5.4 Similarity to Conventional Methods

It was shown in Section 4.2.1 that computing the centroid using an FIR filter in this way is equivalent to performing the convolution of the original function $f(x)$ with some other function $h(x) = -kx$ (see Equation 4.2.4). A common method for locating turning points in time is to take the first order differential of some function and then smooth the result by convolving with a band limited window. For example, if the input function is $f(x)$ and the smoothing function is $g(x)$:

$$y(x) = f'(x) * g(x) \tag{4.5.1}$$

Using a smoothing function of the form $-x^2$, and noting that $f'(x) * g(x) \equiv f(x) * g'(x)$, Equation 4.5.1 may be rewritten as:

$$y(x) = f(x) * -2x \tag{4.5.2}$$

Equation 4.5.2 is now the same form as Equations 4.2.3 - 4.2.5, that is to say that the convolution computed by the centroid filter is comparable to taking a first order difference

and smoothing using a band limited function. Of course the implementation of the centroid filter provided in this Chapter is substantially more efficient than a direct computation of the latter.

4.6 Conclusions

A new method has been described that significantly improves one of the key issues in neural spike sorting, i.e. the difficulty of correctly aligning APs in the time domain in the presence of noise. A novel real time approach for spike alignment based on a *centre of gravity* (centroid) filter has been described that provides an alternative to traditional spike alignment methods and substantially improves resilience to noise and resulting sampling jitter. Validation of the new methods has been achieved by simulation using deterministic models of nerve signals (APs with added noise; both correlated and uncorrelated noise models have been considered). In addition it is shown that the new method lends itself particularly well to hardware realization and a power efficient solution has been described that operates in real time on a single Altera Max V CPLD. The technique has the potential to influence significantly the design of electrophysiological recording systems in the future.

References

[1] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials.," *Network (Bristol, England)*, vol. 9, pp. R53–78, Nov. 1998.

[2] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

[3] S. Narasimhan, X. Wang, and S. Bhunia, "Implantable electronics: emerging design issues and an ultra light-weight security solution.," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2010, pp. 6425–8, Jan. 2010.

[4] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. N. Donaldson, and J. T. Taylor, "An Enhancement to Velocity Selective Discrimination of Neural Recordings: Extraction of Neuronal Firing Rates," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2014, pp. 4111–4114, Aug. 2014.

[5] D. A. Adamos, E. K. Kosmidis, and G. Theophilidis, "Performance evaluation of PCA-based spike sorting algorithms.," *Computer methods and programs in biomedicine*, vol. 91, pp. 232–44, Sept. 2008.

[6] R. Schoonhoven and D. F. Stegeman, "Models and analysis of compound nerve action potentials.," *Critical reviews in biomedical engineering*, vol. 19, pp. 47–111, Jan. 1991.

[7] M. Protter and C. Morrey, "College calculus with analytic geometry," p. 526, 1977.

[8] P. Kwan and C. Clarke, "FPGAs for improved energy efficiency in processor based systems," *Advances in Computer Systems Architecture*, pp. 440–449, 2005.

[9] C. Chow, L. Tsui, P. Leong, W. Luk, and S. Wilton, "Dynamic voltage scaling for commercial FPGAs," *Proceedings. 2005 IEEE International Conference on Field-Programmable Technology, 2005.*, pp. 173–180, 2005.

[10] B. Metcalfe, D. Chew, C. Clarke, N. Donaldson, and J. Taylor, "Fibre-selective discrimination of physiological ENG using velocity selective recording: Report on pilot rat experiments," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2645–2648, IEEE, Aug. 2014.

[11] N. Dalkiliç and F. Pehlivan, "Comparison of fiber diameter distributions deduced by modeling compound action potentials recorded by extracellular and suction techniques.," *The International journal of neuroscience*, vol. 112, pp. 913–30, Aug. 2002.

[12] J. Segundo, J. Vibert, K. Pakdaman, M. Stiber, and O. D. Martinez, "Noise and the neurosciences: a long history, a recent revival and some theory," *Origins: Brain and Self Organization*, p. 299, 1994.

[13] L. M. Ricciardi and L. Sacerdote, "The Ornstein-Uhlenbeck process as a model for neuronal activity," *Biological Cybernetics*, vol. 35, pp. 1–9, Mar. 1979.

[14] J. J. Struijk, "The extracellular potential of a myelinated nerve fiber in an unbounded medium and in nerve cuff models.," *Biophysical journal*, vol. 72, pp. 2457–69, June 1997.

[15] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "A summary of the theory of velocity selective neural recording.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2011, pp. 4649–52, Jan. 2011.

[16] C. T. Clarke, X. Xu, R. Rieger, J. Taylor, and N. Donaldson, "An implanted system for multi-site nerve cuff-based ENG recording using velocity selectivity," *Analog Integrated Circuits and Signal Processing*, vol. 58, pp. 91–104, Nov. 2008.

[17] H. K. Jung, J. H. Choi, and T. Kim, "Solving alignment problems in neural spike sorting using frequency domain PCA," *Neurocomputing*, vol. 69, pp. 975–978, Mar. 2006.

Chapter 5

Velocity Spectral Density

5.1 Extraction of Functional Information - The Need for a New Method

Until recently the principle of *velocity selective recording* (VSR) has only been applied to acute *ex vivo* recordings made from *pig* [1] and *frog* [2] where the nerve fibres have been explanted and selectively recruited using electrical stimulation. In contrast to natural activation, electrical stimulation recruits, approximately simultaneously, every axon within the nerve for which the stimulation current is supra-threshold. Electrical stimulation therefore produces a *compound action potential* (CAP), wherein each axon of a particular diameter fires simultaneously. The *intrinsic velocity spectrum* (IVS) of this type of recording is found by varying an artificially inserted delay and taking the peak of the response as a measure of the number of fibres that were excited at the corresponding velocity (see Equation 3.1.1) [2]. However this approach does not, by itself, result in a method to measure the firing rate for axons of particular velocities, which is of interest when analysing the physiological *electroneurogram* (ENG).

Consider the case where two *action potentials* (APs) of the same amplitude and conduction velocity occur at different times within the same recording, the maximum value observed for the matched delay will be the same as it would be for a single AP. This loss of information precludes the use of the IVS (or in fact any other method that produces a spectrum based only on peak amplitude) for the analysis of firing rates of particular neurons. This is illustrated in Figure 5.1.1 where the IVS has been computed for a time domain recording that contains ten APs each of which are propagating at a conduction velocity of 14 m/s. There were five channels of simulated data and the amplitudes of the APs have been normalised to 1 mV, the amplitude of the peak within the velocity spectrum of Figure 5.1.1 is a function purely of the number of channels and the maximum AP amplitude after the delay-and-add process. For five channels of data with AP amplitudes of 1 mV the peak amplitude within the IVS is 5 mV. The amplitude of the peak is therefore independent of the number of APs and so conveys no information about the

relative firing rates [3].

The firing rate of a particular neuron is key to the transmission of information from one location within the nervous system to another. For example micro-channel electrode structures have been used to record from the teased L6 dorsal root of rat in order to extract information about bladder fullness [4]. In the study the use of micro-dissected rootlets produced a corresponding increase in the *signal-to-noise ratio* (SNR) and so allowed the use of conventional spike sorting techniques. Distinct APs were recorded from the rootlets in response to cutaneous stimulation, bladder filling, and bladder contraction. Offline spike sorting identified three main waveforms from the recordings that correlated to bladder volume, the waveform with the best correlation ($R^2 = 0.74$) varied almost linearly in firing rate from 0 Hz to 15 Hz over the range of bladder volumes $0 - 400$ μL [4]. Once this waveform has been correctly and reliably identified it could be used to control a bladder stimulator device. This method demonstrated that functional information about the bladder could be extracted from a dorsal root, but the electrode configuration required delicate micro-dissection of the fascicles and the spike sorting process depends largely on offline and individual specific processing. If VSR can be modified and extended to produce the firing rates for particular neurons for then it may be possible to construct a real time signal processing system using an easily implantable and stable electrode structure such as a cuff.

This chapter details the modifications to conventional VSR that are necessary to extract firing rates for neurons based on axonal conduction velocity. Several methods are considered that may enable the extraction of spike trains using VSR and one method in-particular (the *velocity spectral density* (VSD)) is analysed in detail and has been applied to simulated data in order to verify and demonstrate the principles required.

5.1.1 RMS-Intrinsic Velocity Spectrum

A proposed method of extending the IVS to include information about the firing rate of particular neurons is to consider the *root mean squared* (RMS) value of each delay-and-add waveform, as opposed to just the maximum value. The basic concept is that the RMS value of each V_d waveform will be a function of the firing rates for particular neurons. Therefore the more frequently a neuron with corresponding conduction velocity (d/t) fires, the larger the RMS signal amplitude will be over a given period of time. The RMS values for each V_d waveform may now be found by modifying Equation 3.1.1 to give:

$$V_{DRMS}[n, dt] = \sqrt{\frac{1}{C}\sum_{i=1}^{C} V_{Bi}[n - (i-1).dt]} \tag{5.1.1}$$

In order to examine the effectiveness of this method a simulation study was performed. A simplified *transmembrane action potentia*l (TMAP) model (Equation 5.1.2) was used to generate artificial data representative of a *multiple electrode cuff* (MEC) 20 cm in length

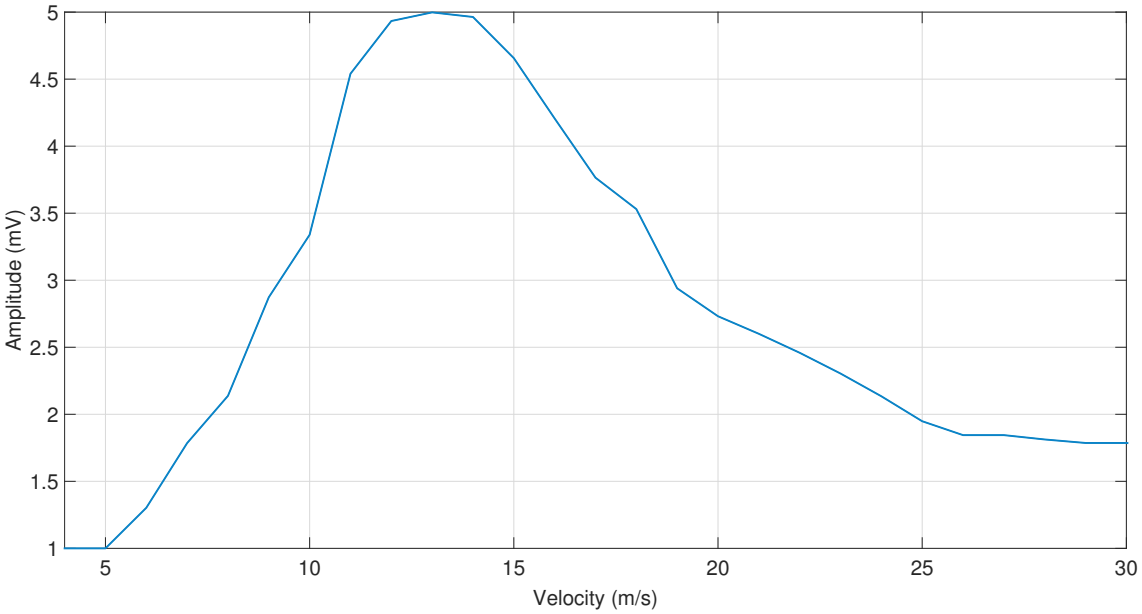


Figure 5.1.1: The IVS for a time domain recording containing ten APs propagating at 14 m/s. There amplitude of the peak is independent of the number of APs and so conveys no information about the relative firing rates.

with eleven equidistant electrodes and a sampling rate of $f_s = 500$ kHz:

$$V_m(t) = At^n e^{-Bt} \tag{5.1.2}$$

Where A is the amplitude, t is time and n and B are free parameters used to control the shape of the AP.

These parameters were chosen to be representative of existing cuff structures and recording systems (*cf* [5, 6]). Nine tripolar signals were formed from a simulated time segment of length 250 ms containing one hundred APs. The APs were randomly selected from two populations of 10 m/s and 15 m/s with equal probabilities and random offsets within the time recording. The random generator was initialised using the Mersenne twister (a commonly used pseudorandom number generator) and designed so that although the time offsets for each AP were random there were an equal number at each velocity.

Figure 5.1.2 illustrates a short time segment from a single channel of the simulated tripolar data, clearly visible are five APs - one of which is classified as a compound AP, where two APs have overlapped in time. Note that the tripolar recording configuration produces a tri-phasic waveform within the time domain. Figure 5.1.3 (left inset) illustrates the IVS for the entire 250 ms recording and clearly shows the two populations. Note however the amplitude variation between the two populations: there were fifty APs at 10 m/s and fifty APs at 15 m/s and so it is desirable for the amplitudes within the velocity spectra to reflect this. Figure 5.1.3 (right inset) shows the RMS velocity spectrum computed by taking the RMS value over the entire 250 ms window (as opposed to the maximum value used for the intrinsic spectrum). The amplitudes at each conduction velocity are now more balanced but the peak velocity has shifted from 15 m/s to 14 m/s.

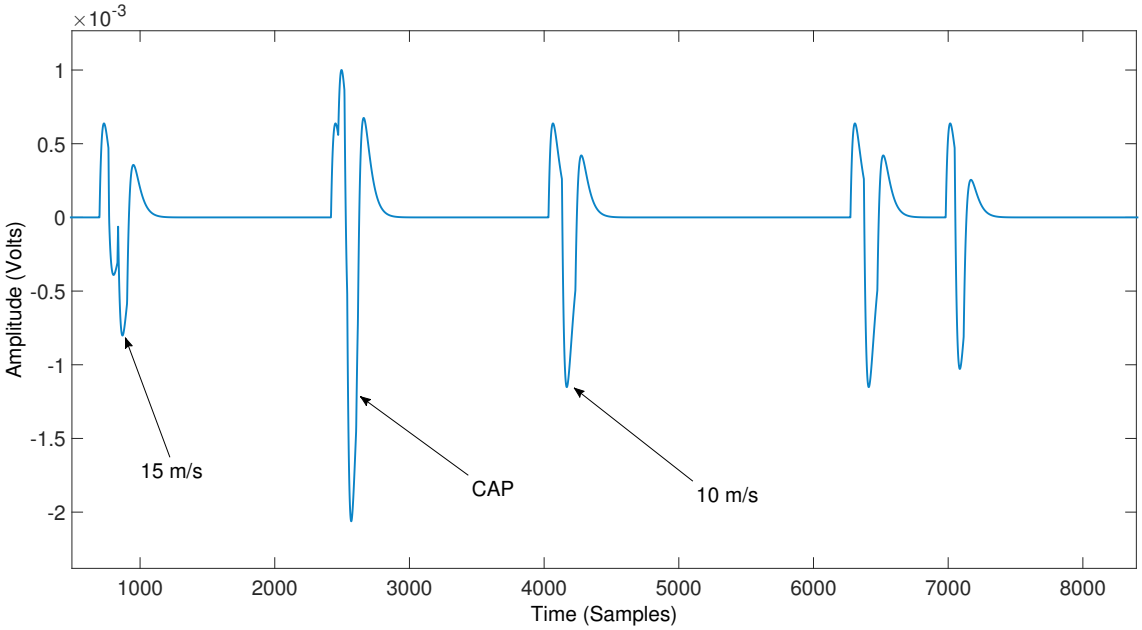


Figure 5.1.2: A segment of the simulated time domain data showing five APs including two CAPs (time overlapping APs).

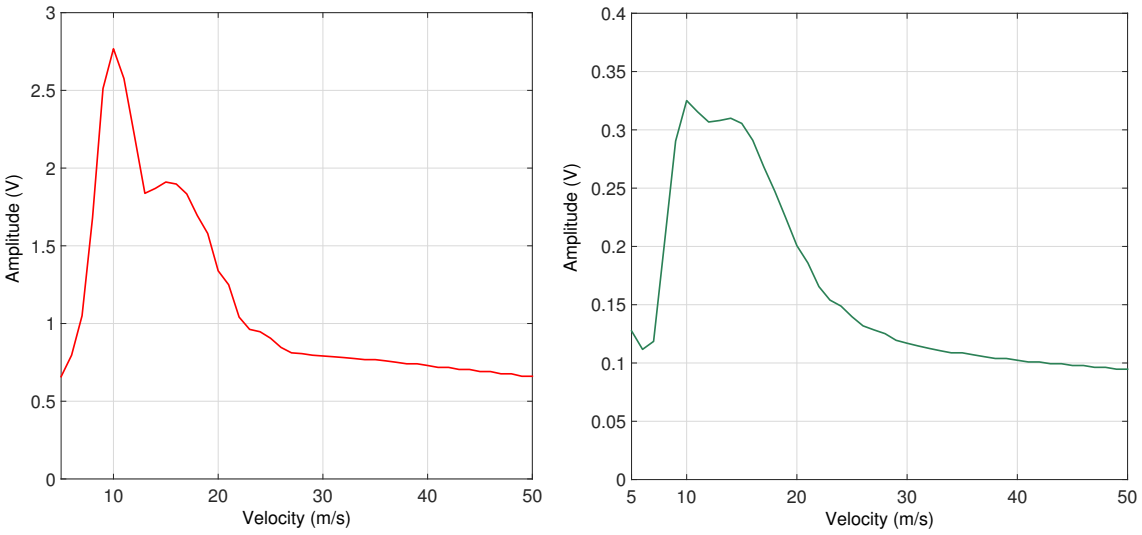


Figure 5.1.3: Left Inset - The IVS for the entire 250 ms recording. **Right Inset -** The RMS velocity spectra for the same 250 ms recording.

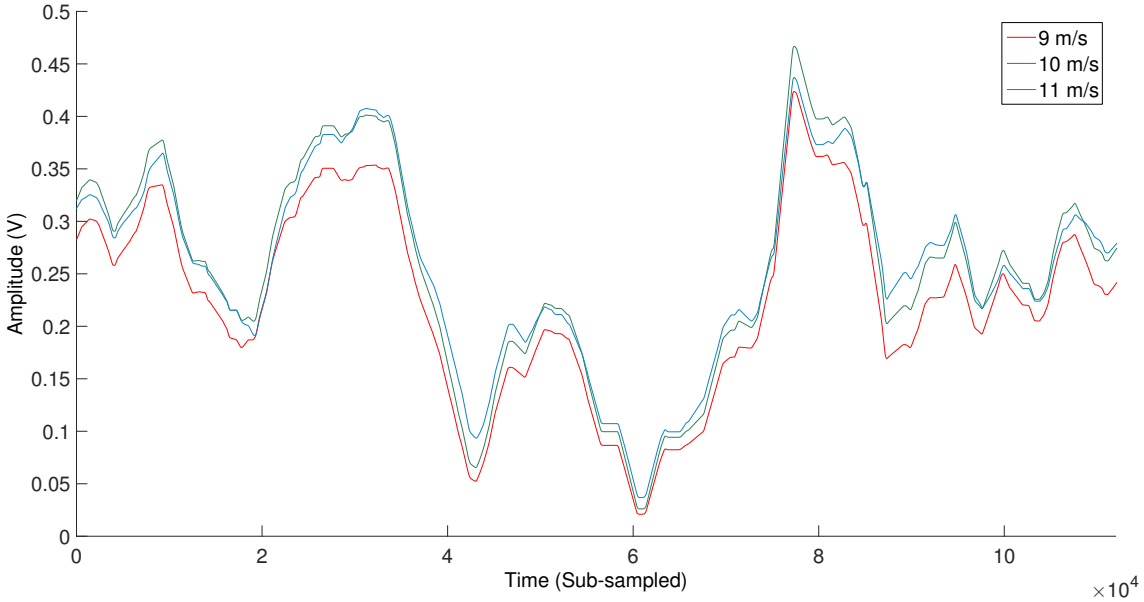


Figure 5.1.4: RMS values computed using a sliding time window of 4 ms (the time taken for an AP of 5 m/s to propagate the length of the MEC) for three V_D waveforms around 15 m/s, the RMS values have been smoothed using a moving average filter 20 ms long.

In a recording system designed for physiological ENG it is a requirement to record the firing rates of neurons over a fixed time period, which may range from a few milliseconds to tens of seconds depending on the particular neuron. The RMS spectra may be calculated from a sliding time window, whose length would be chosen based on the desired integration time. Consider the plots shown in Figures 5.1.4 and 5.1.5, which represent the RMS values for three neighbouring velocities (V_{DRMS}). The RMS spectra were computed using a sliding time window that was 4 ms long and the values have been smoothed with a three point moving average (box car) filter that is 20 ms long. A moving average filter is required to smooth over periods where there may not have been any APs within the sliding time window. Figures 5.1.4 and 5.1.5 illustrate the limiting constraint of the RMS approach, the ratio in amplitudes between neighbouring velocities (i.e. the *velocity selectivity*) even in this noise free simulation is relatively poor and there is little differentiation between the matched and the mismatched velocities. While the use of an RMS spectrum does provide a measure of the specific firing rates it is inherently limited by the intrinsic velocity selectivity of the recording system. Further to this there are multiple free parameters that must be selected by the designer in advance, namely the length of time window over which to take an RMS value and also the length of the moving average filter.

A more robust method has been developed that does not require the use of a sliding time window or any prior knowledge about the firing rate of the neurons, and so avoids these issues. The new method by which both conduction velocity and neuronal firing rates can be extracted is termed the method of *velocity spectral density* (VSD) and is described in the following sections.

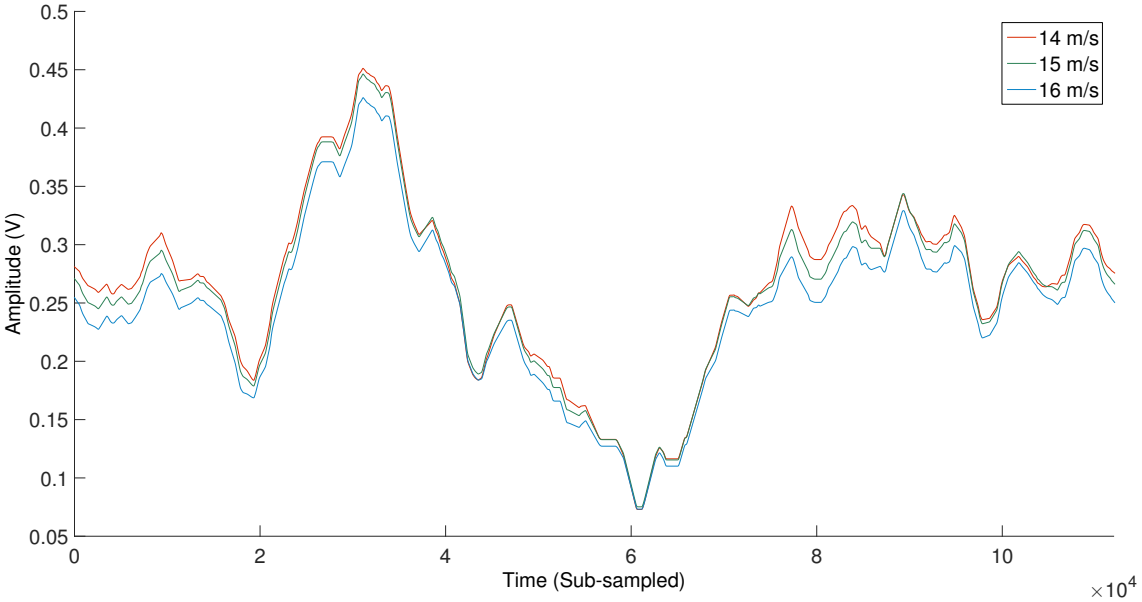


Figure 5.1.5: RMS values computed using a sliding time window of 4 ms (the time taken for an AP of 5 m/s to propagate the length of the MEC) for three V_D waveforms around 15 m/s, the RMS values have been smoothed using a moving average filter 20 ms long.

5.2 Principle of Velocity Spectral Density

5.2.1 Generation of Velocity Specific Waveforms

VSD can be considered to be a natural extension of traditional delay-and-add based methods and it attempts to quantify the activity within prescribed bands of conduction velocity as discrete firing rates, or the number of APs occurring per second. The first stage of the VSD process is to perform the delay-and-add process on a set of recordings made from a multiple electrode system. Equation 5.2.1 is the conventional form of delay-and-add and was described in detail in Chapter 3:

$$V_D[n, dt] = \sum_{i=1}^{C} V_{Bi}[n - (i-1).dt] \tag{5.2.1}$$

The artificially inserted delays, which are multiples of dt, can be selected based on the required velocity range and resolution. For example a velocity range of 10 m/s – 50 m/s with an electrode spacing of 1 mm requires delay, or dt, values in the range 20 μs – 100 μs. If the resolution is 1 m/s then there will be $m = 41$ V_D waveforms, each representative of a different artificial delay dt. An example is shown in Figure 5.2.1 where five channels of raw data containing a single biphasic AP propagating at 15 m/s have been summed together with values of dt corresponding to conduction velocities of 13, 14, 15, 16 & 17 m/s from left to right respectively. In this noise free situation it is clear to see that the peak amplitudes (in both the positive and negative phase of the AP) occur at the matched velocity of 15 m/s. The data shown in Figure 5.2.1 were generated using the TMAP model of Equation 5.1.2 using a simulated sample rate of 500 kS/s, an effective electrode spacing

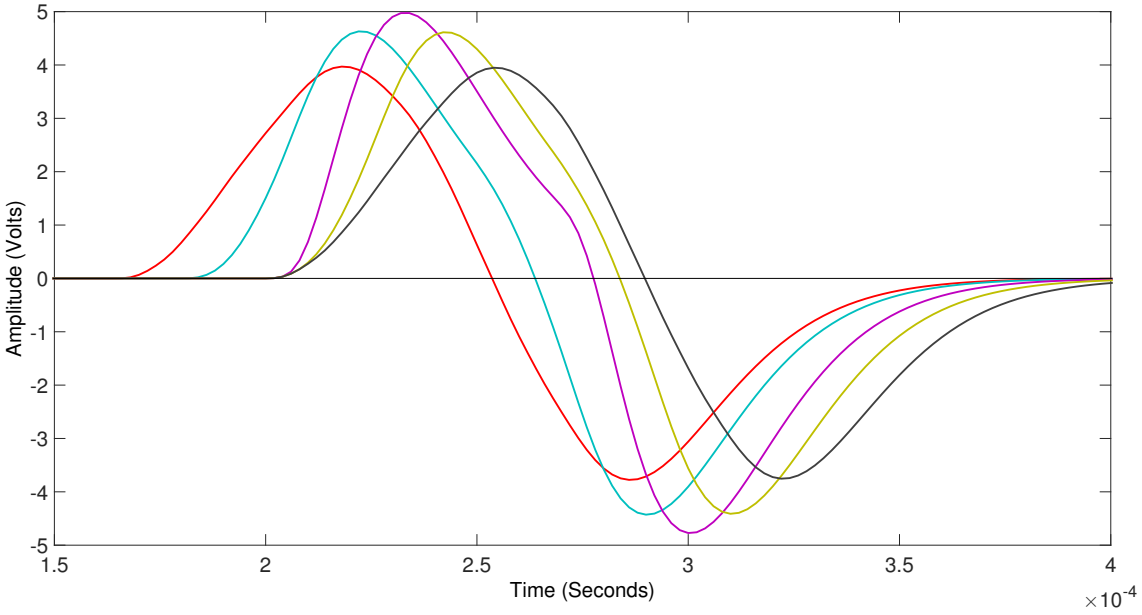


Figure 5.2.1: Typical V_D waveforms for a single AP. Five channels of data (containing a single AP propagating at 15 m/s) have been delayed and summed for values of delays corresponding to velocities of 13 m/s − 17 m/s respectively. It is clear to see the effects on constructive superposition at the matched velocity, the largest positive and negative peak occurs at the delay corresponding to the matched velocity.

of 1 mm and a total of ten bipolar recording channels representative of a short MEC [7]. The relationship between adjacent values of V_D is key to identifying the matched velocity and thus the most likely conduction velocity for the observed AP. In the simple noise free case shown in Figure 5.2.1 it is possible to observe by eye which of the V_D waveforms corresponds to the matched velocity as it possesses the largest peak amplitude. It is possible, however, to define an automatic and on-line process that can identify the matched velocity based on the amplitude relationship between the different V_D waveforms and does not require either human interaction or any prior knowledge about the neuron.

The delay-and-add process can be loosely compared to cross-correlation where the multiplication within the correlation process has been replaced with an addition. The relationship shown in Figure 5.2.1 only occurs when a particular characteristic (in this case the positive phase of the AP) correlates across channel with a fixed and uniform delay. It is possible therefore to identify an AP based purely on the presence of an interchannel correlation - this is a substantial advantage in artefact rejection when compared to single channel recordings. Each V_D waveform is passed through a filter that detects the *centroid* of each AP [3, 8]. The operation of this filter is described in detail within Chapter 4, however a brief overview will be provided here. The *centroid filter* is implemented as a linear *finite impulse response* (FIR) filter with impulse response given by $h[n]$:

$$h[n] = \frac{-2}{N}n + 1 \tag{5.2.2}$$

This is a linear function of gradient $-2/N$ where N is the width of the filter is terms of samples and n is the current index of the discrete-time samples. The function $h[n]$

varies in amplitude from +1 to -1 where N is chosen to be at least as wide as a single AP in the time domain. Since in practice the AP is neither regular nor symmetric the centroid represents a more robust method for locating the midpoint of the AP than taking the maximum value. The maximum value is also a single value measure and is easily perturbed by additive noise. The centroid can be considered as the geometric centre of any two dimensional region, in this case the area under the AP as bounded by the x axis. It is necessary to separate the positive and negative phases of the AP before locating the centroid, and this was achieved via half wave rectification of the signal. It is assumed throughout that there is zero DC offset in the data (in practice this can be achieved by band-pass filtering the data). Computing the centroid considers the contribution from every sample as opposed to the single samples used in peak detection and so it is more robust against noise and interference. The filter output can now be described by $y[n, dt]$ which is the convolution of $h[n]$ with $V_D[n, dt]$:

$$y[n, dt] = \sum_{k=-\infty}^{\infty} \left[\left(\frac{-2}{N}(n-k) + 1 \right) . \left(\sum_{i=1}^{C} V_{Bi}(n - (i-1).dt \right) \right] \tag{5.2.3}$$

$y[n, dt]$ passes through zero at a point that corresponds to the centroid of each AP time shifted by the group delay of the filter T_g, which, due to negative symmetry in the impulse response, is given by:

$$T_g = \frac{(N-1)T}{2} \tag{5.2.4}$$

Where T is the sampling interval and N is the filter order. Whenever a zero-crossing is detected, the instantaneous value of V_D is held:

$$V_D[n, dt] = \begin{cases} y[n, dt], & y[n-1] > 0 > y[n+1] \\ 0, & otherwise \end{cases} \tag{5.2.5}$$

The process is illustrated by the example shown in Figure 5.2.2 where a set of V_D waveforms corresponding to conduction velocities of 13 m/s - 17 m/s from left to right have been gated at their relative centroids, the group delay of the filter has been corrected by using a simple delay line element. Note that in this simple case only the positive phase of the AP has been considered although the process could be easily extended to generate separate gated pulses for the negative phase of the AP, and recent results with electrically stimulated nerves from *frog* indicate that the apparent conduction velocity of the AP may differ depending on the polarity used for the measurement [2]. This gating process could also be extended to include propagation in both *afferent* and *efferent* directions, thereby generating a total of four gated pulse trains that describe the interchannel correlation of APs travelling in both directions along a nerve.

A simple detection algorithm can now be applied that examines each velocity response $V_D[n, dt]$ for the criteria $V_{D-1} < V_D > V_{D+1}$. The held value of $V_D[n, dt]$ is compared to a noise threshold, calculated from measurements of the input-referred noise floor, and then compared across the m data streams. As shown in Figure 5.2.2 the gated centroid of

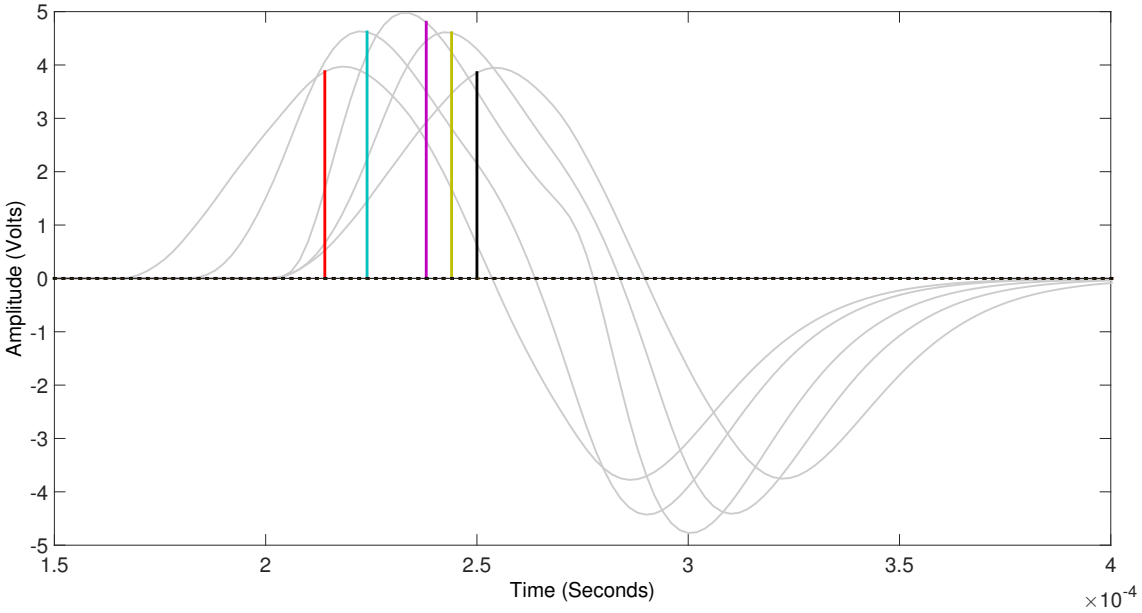


Figure 5.2.2: Using the output from each centroid filter the V_D data steams may now be gated (or held) for single sample values at the centroid of each AP. The result spikes (which are shown here superimposed on the original V_D waveforms) may now be inspected for the detection criteria. In this case it is clear that the most likely conduction velocity occurs for a value of V_D corresponding to 15 m/s, the range of velocities is 13 m/s − 17 m/s from left to right respectively.

each $V_D[n, dt]$ waveform is offset in time by a multiple of the sampling interval and so the examination of each response is not trivial. In order to compare the amplitudes the value of each gated waveform is held in memory until the next AP occurs, i.e. when the gated value of $V_D[n, dt]$ is next non-zero. It is then straightforward to compare the values stored in memory. If this criterion is met then an AP has been detected with a peak conduction delay (and thus velocity) of dt.

Once the detection criterion has been met a marker may be inserted into a new time record that indicates at what time the AP was detected. Each conduction velocity has a separate time record, this is akin to traditional spike sorting methods where markers are inserted into time records based on AP morphology. Figure 5.2.3 illustrates the spike trains for three hypothetical neurons, in morphological spike sorting the three neurons are classified based on the amplitude and width of the AP, in VSD based analysis the three neurons are classified based on conduction velocity of the AP along the recording structure. Additionally it is possible to construct histograms that illustrate the comparative firing rates between different bands of conduction velocity. These are readily computed by simply counting the number of markers within each spike train over a fixed integration time (typically one second). A modified VSR process is illustrated in Figure 5.2.4, the conventional delay-and-add processes has been augmented with the centroid filtering and gating process as well as the detection algorithm. Note that V_{D-1} and V_{D+1} must be available for comparison, and so it is necessary to compute V_D waveforms corresponding

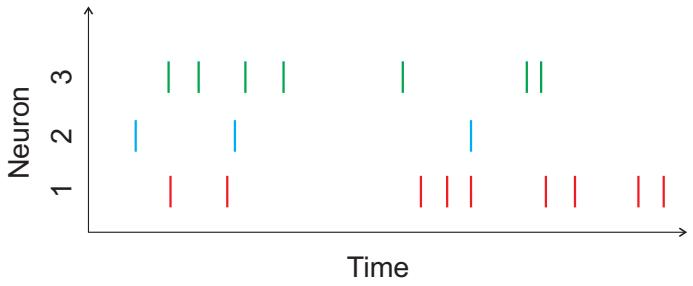


Figure 5.2.3: Example spike train illustrating the relative firing times for three hypothetical neurons, the classification may be performed based on AP morphology (spike sorting) or on conduction velocity (VSD/VSR).

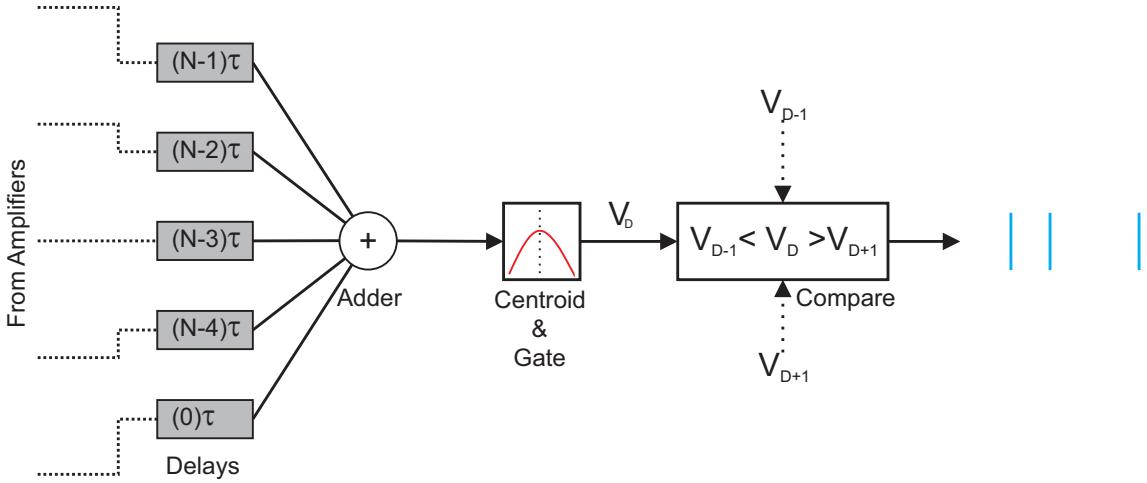


Figure 5.2.4: A modified VSR processor that has been extended to produce velocity specific spike trains using the VSD method.

to at least three conduction velocities.

5.3 Results for Simulated Data

In order to validate the VSD method a set of simulations were performed where theoretical models for the TMAP were used to generate artificial time domain recordings that are representative of a myelinated nerve within a MEC. The TMAP model of Equation 5.1.2 was used throughout the simulations with the following parameters: $A = 2.6 \times 10^1$ $B = 1.5 \times 10^4$ and $n = 1$ [9]. These parameters are representative of a large myelinated axon. The sampling rate was 500 kS/s and the total length of the time domain simulation was 100 ms. Five channels of bipolar data were simulated for a MEC of length $5 \times 10^{-3} m$ with a total of six equidistant electrodes. Ten APs were generated with random conduction velocities within the range 5 m/s – 15 m/s and arbitrary but non overlapping start times. Table 5.3.1 details the simulated APs. VSD analysis was applied for a range of dt values corresponding to a conduction velocity range of 5 m/s – 15 m/s with a resolution of 1 m/s and the noise threshold was set to be 1 V. Figure 5.3.1 shows the resulting spike trains

Table 5.3.1: Simulated AP conduction velocities, position within the time domain and peak amplitudes.

AP ID	Conduction Velocity (m/s)	Time Offset (ms)	Peak Amplitude (mV)
1	10	8	0.06
2	12	16	0.08
3	7	22	0.03
4	8	50	0.04
5	8	55	0.04
6	7	62	0.03
7	12	70	0.08
8	13	76	0.10
9	10	84	0.06
10	7	90	0.03

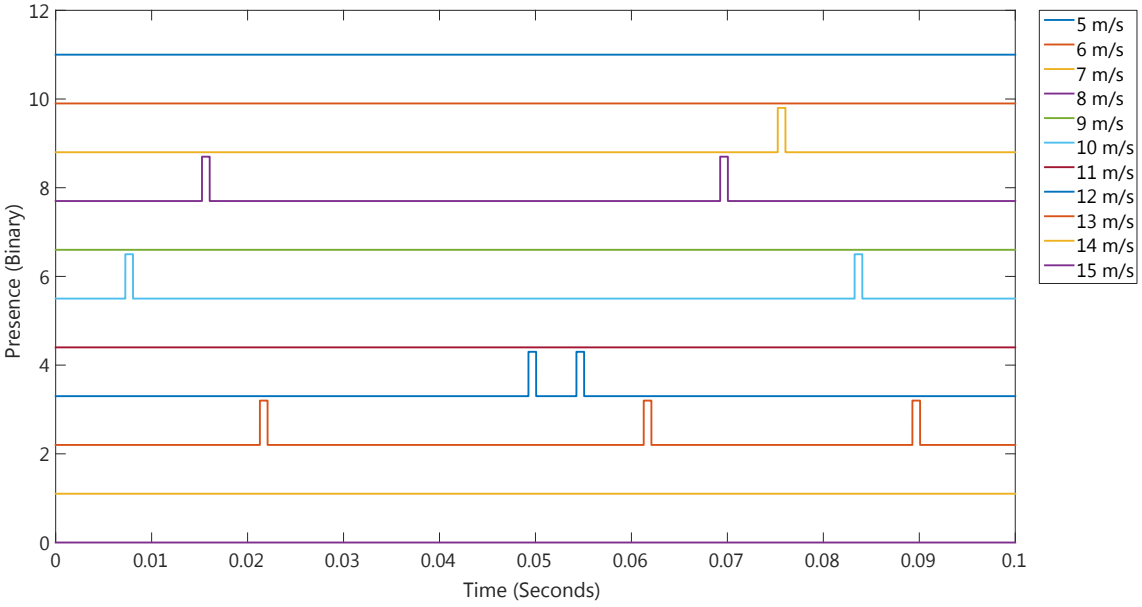


Figure 5.3.1: VSD generated spike trains in the range of 5 m/s−15 m/s extracted from noise free simulated data, an arbitrary pulse of unit amplitude and width 400 samples is generated when an AP of the corresponding conduction velocity is detected.

produced using this process, the width of each of the artificially generated markers is a free parameter and in this instance was set to be 400 samples. Each of the ten APs was correctly identified both in terms of conduction velocity and location within the time recording; the VSD produced accurate spike trains with minimal complexity on top of traditional delay-and-add processing.

5.3.1 Results with Noise

Fundamentally the VSD performs an amplitude comparison between three adjacent conduction velocities after the delay-and-add process. It was shown in Chapter 3 that the delay-and-add process provides an increase in SNR of approximately $\sqrt{N}$, where N is the number of recording channels. Therefore if the SNR in the time domain recordings is below 0 dB it may still be possible to apply the VSD if there are sufficient recording chan-

nels such that the post delay-and-add SNR is greater than 0 dB. Further tests performed with *additive white Gaussian noise* (AWGN) revealed that there are a number of possible failure modes for the VSD as the post delay-and-add SNR is reduced towards zero. The primary failure mode is the lack of any correlated features above the noise floor, leading to false negatives, or the failure to identify particular APs. The secondary failure mode occurs when the noise threshold is set too low and leads to false positives, or the identification of noise processes as APs. The third and final failure mode is the misidentification of the peak velocity, this can occur when the largest amplitude response (largest gated V_D) is not the correct velocity. Simulations performed with AWGN have shown that for nonoverlapping APs the VSD correctly identifies the conduction velocity and AP location for SNR values above 1 dB after the delay-and-add process. Below this value of SNR the VSD fails entirely, that is to say that there is very little degradation in performance until the method is unable to identify any APs because the noise threshold required excludes the entire waveform. For overlapping APs the situation is more complex, if the recording array is long enough then separation of APs with different conduction velocities will occur naturally over the length of the array. For shorter arrays the APs will remain overlapping and the VSD will produce only a single output that is representative of the average conduction velocity for the two APs. Within the V_D waveforms experiments have shown that the peak gated values must be at least 100 samples apart (corresponding in this work to about half the width of an AP) in order for the VSD to correctly identify both APs. A more detailed analysis of the noise performance of the centroid filter under both correlated and uncorrelated noise is given in Chapter 4.

5.4 Conclusions

To date VSR has only been applied successfully to the analysis of acute recordings of the electrically evoked ENG, a situation in which a single CAP is recruited for each population (axons with the same conduction velocity). In this type of recording the IVS is an appropriate tool and gives a good indication of the velocity content within the recording and thus the probable diameters of the axons, however the fundamental limitation of the IVS is that it provides no information about the *firing rates* of each neuron. It has been well documented that it is the firing rate of the neuron that encodes *somatosensory* information within the nervous system, and in part it is this limitation that has prevented the application of VSR based methods to naturally occurring or physiological ENG. A proposed method of extending VSR using RMS spectra has been examined using a simulation study and it has been shown that this method requires a very high level of velocity selectivity; with currently available electrode structures the practical selectivity is insufficient for the RMS spectra to give an accurate indication of firing rates. Furthermore the RMS spectral method requires a sliding time window and smoothing function that must be carefully designed based on prior knowledge about the expected firing rate of the axon. The VSD method that has been described in this Chapter is a natural extension

to VSR and does not require any such prior information; rather it takes advantage of the naturally-occurring interchannel correlation in order to extract spike trains, and consequently neuronal firing rates, based on the conduction velocity of each AP. A validation of the VSD method has been provided using simulated data and in the following chapters the method is applied to naturally occurring physiological ENG, for the first time providing a detailed analysis of both conduction velocity and firing rate in a completely on-line and unsupervised process.

References

[1] M. Schuettler, V. Seetohul, N. J. M. Rijkhoff, F. V. Moeller, N. Donaldson, and J. Taylor, "Fibre-selective recording from peripheral nerves using a multiple-contact cuff: Report on pilot pig experiments.," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 3103–6, Aug. 2011.

[2] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[3] B. Metcalfe, D. Chew, C. Clarke, N. Donaldson, and J. Taylor, "Fibre-selective discrimination of physiological ENG using velocity selective recording: Report on pilot rat experiments," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2645–2648, IEEE, Aug. 2014.

[4] D. J. Chew, L. Zhu, E. Delivopoulos, I. R. Minev, K. M. Musick, C. A. Mosse, M. Craggs, N. Donaldson, S. P. Lacour, S. B. McMahon, and J. W. Fawcett, "A microchannel neuroprosthesis for bladder control after spinal cord injury in rat.," *Science translational medicine*, vol. 5, 2013.

[5] M. Schuettler, V. Seetohul, J. Taylor, and N. Donaldson, "Velocity-selective recording from frog nerve using a multi-contact cuff electrode.," *Conference Proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 2962–5, Jan. 2006.

[6] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. N. Donaldson, and J. T. Taylor, "An Enhancement to Velocity Selective Discrimination of Neural Recordings: Extraction of Neuronal Firing Rates," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2014, pp. 4111–4114, Aug. 2014.

[7] G. E. Loeb and R. A. Peck, "Cuff electrodes for chronic stimulation and recording of peripheral nerve activity.," *Journal of neuroscience methods*, vol. 64, pp. 95–103, Jan. 1996.

[8] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. d. N. Donaldson, and J. T. Taylor, "A new method for spike extraction using velocity selective recording demonstrated with physiological ENG in Rat.," *Journal of neuroscience methods*, vol. 251, pp. 47–55, Aug. 2015.

[9] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

Chapter 6

Pilot Experiments in *Rat*

6.1 Motivations and Overview

In spite of the many potential applications of *velocity selective recording* (VSR), to date it has only been demonstrated with electrically evoked *electroneurogram* (ENG), i.e. *compound action potentials* (CAPs) in *pig* and *frog* [1, 2]. Furthermore, there have as yet been no successful attempts to apply VSR to naturally occurring neural signals.

There are several significant differences between the requirements of recording electrically evoked and natural (physiological) ENG that complicate the process of recording the latter. These differences include much smaller signal amplitudes (typically, using cuffs, $1\ \mu\text{V} - 10\ \mu\text{V}$, as opposed to around $100\ \mu\text{V}$ for CAPs) and the need to determine the rate of neural firing in a particular velocity band rather than the relative amplitudes of activity between bands, which is generally the case for electrically evoked potentials.

This chapter presents the first experimental validation of the *velocity spectral density* (VSD) method that was presented in Chapter 5, as well as the first ever application of velocity selective methods to naturally evoked (physiological) ENG. Experiments were performed to obtain physiological ENG from a strand of an intact nerve within a dorsal sensory root of *rat*. Both the VSR and the VSD methods were employed to provide an in-depth analysis of the recorded data and to overcome the difficulties of recording physiological ENG, thereby providing a method that allows neuronal firing rates in specified velocity bands to be identified in real time [3]. The surgical aspect of this chapter was kindly performed by Dr. Daniel Chew, of the Jon Van Geest Centre for Brain Repair at the University of Cambridge, the electrodes were fabricated and the experiments designed with the aid of Prof. Nick Donaldson at the Implanted Devices Group of University College London.

The recordings were made with the nerve resting on *hook* electrodes whilst immersed in nonconductive paraffin oil. This was convenient for the experiment but the signals so obtained were similar to the outputs from a row of electrodes in a micro-channel struc-

ture. Both arrangements greatly increase the amplitude of extracellular potentials (V_{ex}) compared to a nerve of diameter 1 mm or greater placed in a cuff, enabling individual *action potentials* (APs) to be distinguished and counted: this is essential for the method and also allows validation of the ENGs by inspection. The ten electrodes were connected in separate pairs to form five bipolar recording channels before amplification, band-pass filtering and digitisation.

In order to validate the VSD method as applied to physiological ENG a number of APs were visually identified in the time recording. The conduction velocity of each AP was calculated by hand using the delay between the peak values from one channel to another. The VSD process was then applied to the same data set and the resulting conduction velocities compared to the values found by inspection with very close agreement. The use of hook electrodes in conjunction with the insulating oil was key in providing the high level of *signal-to-noise ratio* (SNR) required for the manual comparison.

Two sets of experiments were performed using different constructions of the hook electrodes, the latter using flexible hooks that more readily conformed to the shape of the nerve. The results presented are predominantly from the second set of experiments unless otherwise indicated. The experimental setup did not differ otherwise from one experiment to the other. Modulation of the neural signals was elicited by both manual and electrical stimulation of the derma at a point known to be innervated by the L5 root and VSD was used successfully to detect the onset of stimulation in the physiological ENG.

There is, of course, a long history of analysing neural recordings from microelectrodes in *brain*. These methods normally identify APs by their characteristic morphology. This approach, referred to generically as *spike sorting* is generally not done in real time and the methods often use substantial computing power [4]. By comparison, the proposed VSD method, in common with other VSR-based approaches, can operate in real time and is relatively economical in terms of computational effort. These features are important in certain neuroprosthetic devices such as the "Bioelectronic Medicines" currently being advocated by GlaxoSmithKline [5]. This is because the devices must be small and low-powered and the firing rates of neurons that serve different functions must be calculated without significant computational delays.

The validation of VSR using physiological recordings of the ENG is a vital step towards creating an implantable device suitable for neuroprosthetics. Methods developed and validated using hook electrodes may be extended to chronically implantable electrode structures such as cuff or micro-channel electrodes.

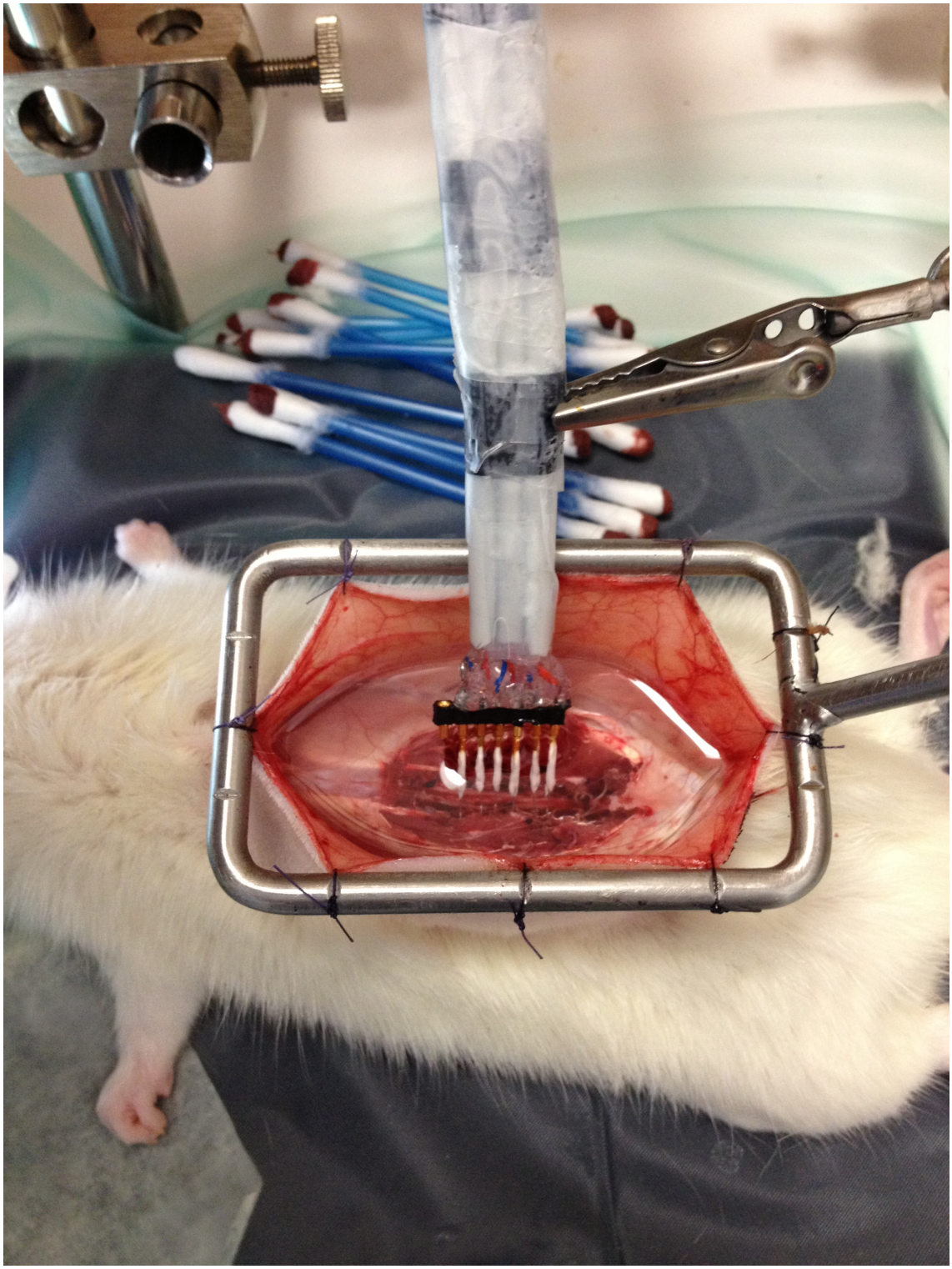


Figure 6.2.1: Typical experimental setup for recording. Clearly shown are six hook electrodes attached to the L5 dorsal rootlet of an adult *rat*, the rootlet and electrodes are submerged in paraffin oil and the electrode assembly and dermis are held in place with a clamp stand and sutures.

6.2 Experiment Setup

6.2.1 Animal Preparation

All animal procedures were performed in accordance with the *United Kingdom Animal (Scientific Procedures) Act of 1986*. In the experiments an adult female Sprague Dawley *rat* (weight approximately 250 grams) was anaesthetised with 1.5 g/kg urethane (sourced from Sigma-Aldrich) administered by the *intraperitoneal* route. The dorsal spinal cord was exposed via a laminectomy of three of the lumbar spinal vertebrae and the animal was placed on an electronic heating mat maintained at 27 degrees Celsius. The dorsal skin was sutured to an overhanging rectangular bar as illustrated in Figure 6.2.1, creating a contained pool into which nonconductive paraffin oil was poured. The *dura* was incised to expose the dorsal roots. The nonconductive oil was used to prevent the roots from dehydrating and also to provide electrical insulation from one hook to another; a thin layer of extracellular fluid remained close to the nerve and provided a current path between the nodes of Ranvier.

The left fifth lumbar dorsal root (L5) was micro-dissected into fine rootlets (fascicles) with fine glass pulled pipettes, as in the method described previously [6–8].

The dorsal root was chosen for a number of reasons. Firstly it is long enough to fit into the multiple electrode array, secondly it is amenable to the micro-dissection technique, and finally it is exclusively sensory, containing only afferent APs. During the course of the experiments, modulation of the neural signals was elicited by stimulating the L5 dermatome both manually (via direct cutaneous touch) and electrically (via bipolar pin electrodes).

One fascicle, approximately 100 μm in diameter, was placed over all of the hooks in the array. The electrodes were arranged in a shared bipolar configuration whereby five channels were recorded from ten independent but commonly-referenced amplifiers in a multiple electrode structure: a simplified recording diagram is illustrated in Figure 6.2.2 [9]. The amplifiers were connected to the recording set-up (Digitimer, UK) and the animal was suitably electrically grounded to the recording equipment and the surrounding Faraday cage. The ten hooks were connected to five unity-gain head-stages (Neurolog NL100) numbered in the *orthodromic* direction of propagation, i.e starting from the soma. Following each head-stage, the signal was preamplified 1000 times using an AC-coupled amplifier (Neurolog NL104A), fed through a 50 Hz interference eliminator (Humbug, Quest Scientific, Canada) and amplified a further 10 times (Neurolog NL106) before band-pass filtering between 300 Hz $-$ 5,000 Hz. The total system gain was 80 dB in the passband.

Table 6.2.1: Specifications of the high speed ADCs used throughout the experiments, the sampling rate was 500 kS/s.

Property	**Value**
Gain Drift	6 ppm/C
Offset Drift	29 μV/C
CMRR ($f = 60$ **Hz**)	100 dB
-3 **dB Bandwidth**	$>$ 500 kHz
Input Impedance	$>$ 1 GΩ
Noise	0.75 LSB
THD (20 $\mathbf{V}_{pp}$ **at** 10 **kHz**)	-85 dB
Crosstalk (20 $\mathbf{V}_{pp}$ **at** 1 **kHz**)	-100 dB

6.2.2 Electrode and Amplifier Configuration

The recording electrode array was fabricated on site, consisting of ten hooks joined to an insulating bar that was supported by a clamp stand. Each hook was formed from 0.2 mm diameter tungsten wire fed through a polyurethane tube of 0.4 mm (internal) diameter. Each wire was fixed in its tube by cyanoacrylate adhesive. The hooks were formed at one end by winding around a cylinder of 4 mm diameter; the shape of each hook was then maintained by the rigidity of the wire. The total length of the electrode array was 5 mm. In previous experiments it was found that it was difficult to maintain a good electrical contact between the nerve and all ten of the hooks. The reason for this was that the nerve was not transected, it was under longitudinal tension and rather than resting on the hooks it was somewhat pulled over them. The use of a thinner wire (0.2 mm) gave each hook some flexibility and allowed the profile of the recording array to bend into a bow shape, resulting in a significantly more reliable electrical contact between the nerve and the hooks.

In addition to the array of recording electrodes, a further pair of pin electrodes was applied to the fifth lumbar (L5) dermatome for electrical stimulation of the skin. The purpose of this was to test the recording system at the start of the experiment and to show the CAP from the cutaneous afferents that will include the touch receptor fibres. The analysis of the CAP using the *intrinsic velocity spectrum* (IVS) provides an estimation of the fibre diameters within the nerve, and thus the expected range of conduction velocities. The pin electrodes were connected to a Neurolog NL200A stimulus isolator driven from an NL301 pulse generator. Square wave stimulation pulses were applied with varying currents (0.3 mA $-$ 4 mA) but with a constant width of 100 μs. The recording equipment was started by means of a synchronisation pulse that was triggered 100 μs before the application of the electrical stimulus.

The amplified and filtered signal was passed to a set of high speed successive-approximation *analogue to digital converters* (ADCs) (NI9222 mounted in cDAQ-9178 by National Instruments, Austin, TX, USA) providing simultaneous sampling of all five bipolar record-

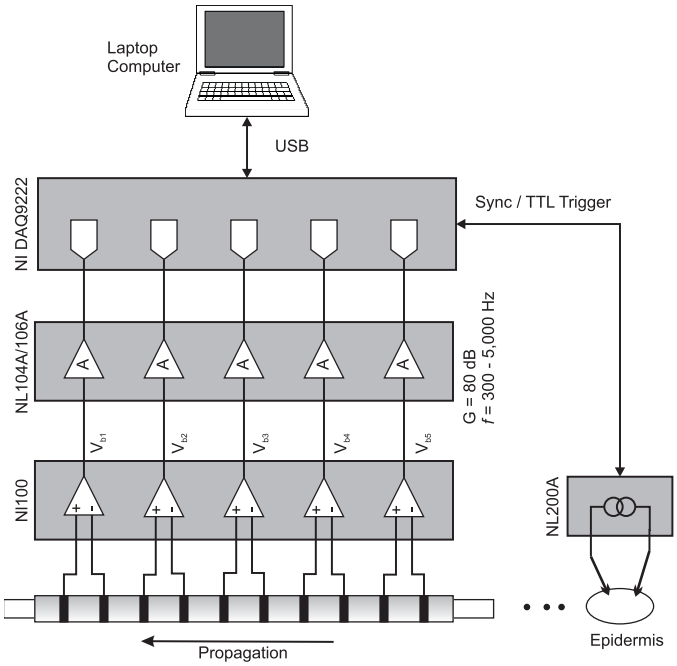


Figure 6.2.2: Recording setup used to acquire physiological ENG. A bank of digital to analogue converters was used to digitise five channels of data simultaneously. The stimulation was controlled by the laptop computer via the data converters using a *transistor-transistor logic* (TTL) pulse. The stimulation electrodes were inserted into the dermis distally from the recording electrodes.

ings with 16 bit resolution. The specifications of the converters are provided in Table 6.2.1.

The converters were connected to a battery operated portable laptop computer running LabView 2010 that logged each channel into a set of data files for offline analysis, no processing was performed on the data during the experiment. Battery operation is preferable as it significantly reduces the likelihood of earth loops within the recording setup, it also avoids the use of the large switch mode power supply unit that produced significant interference to the recordings. Each data file contained a time record and five channels of raw data corresponding to each amplifier output. Online visual verification was provided by connecting a set of oscilloscope channels at the converter inputs, which provided a means to examine each recording in real time for conspicuous artefacts or noise.

6.3 Signal Processing

Offline data analysis was performed using MATLAB R2012b (The MathWorks, Natick, MA, USA). The raw data was imported into MATLAB using a custom file handler designed to automatically extract the individual channels from within the proprietary LabVIEW file format. Each of the five bipolar signals was sampled at a rate of 500 kS/s with sampling occurring synchronously across the 5 channels. Ten consecutive recordings of 250 ms were made. The smallest possible delay was therefore 2 μs corresponding to a maximum velocity

Table 6.4.1: Stimulation currents and the corresponding peak conduction velocities measured from the IVS.

Stimulation Current (mA)	**0.3**	**0.5**	**1**	**2**	**4**
Peak Populations (m/s)	10	9, 14	10	10, 14	7, 14

of 1 mm/2 μs = 500 m/s and allows for a velocity resolution of 0.2 m/s at a velocity of 10 m/s. Each bipolar signal was time-shifted against the others by multiples of dt and then summed to form a single signal. Where the required delay was not a multiple of dt, each bipolar signal was time shifted by alternating values of dt such that the average delay over all five channels was the desired value. The delay-and-add based IVS was used to extract the basic velocity information from the raw data, each recording lasting 250 ms in this case [10, 11]. Further analysis was then performed using VSD methods to create histograms and spike trains for individual velocity bands, the noise threshold used throughout the process was calculated from the input-referred noise floor.

6.4 Results

6.4.1 Electrical Stimulation

Electrical stimulation was applied to the L5 dermatome to test the recording system and record the CAP from the cutaneous afferents. The stimulation waveform was a fixed width (100 μs) square pulse of variable amplitude and Figure 6.4.1 shows the resulting time domain response, the IVS and the V_D waveforms for the peak stimulation current of 4 mA (of length 10 ms). The location of the peaks within the IVS for each value of stimulation current is given in Table 6.4.1.

At lower stimulation currents the peak velocities are at approximately 10 m/s and 14 m/s and as the stimulation current is increased to 4 mA, populations at lower velocities (7 m/s) are recruited in accordance with the principle of *inverse recruitment* [12]. Note that during stimulation with 1 mA and 4 mA the selectivity of the system was not high enough to isolate the populations at 14 m/s and 10 m/s respectively. Figure 6.4.1 illustrates this principle, the peak at 10 m/s has been masked by the larger amplitude signal occurring at 14 m/s - this is why the peak values in Table 6.4.1 contradict the principle of inverse recruitment.

6.4.2 Naturally Evoked (Physiological ENG): Resting State

Recordings were made using the same experimental setup as for electrically evoked stimulation while the animal was in a resting state, i.e. with no external stimulation applied. Ten recordings of duration 250 ms were made at 30 second intervals. The input-referred noise floor was measured during a 5 ms period of no observable neural activity and ranged

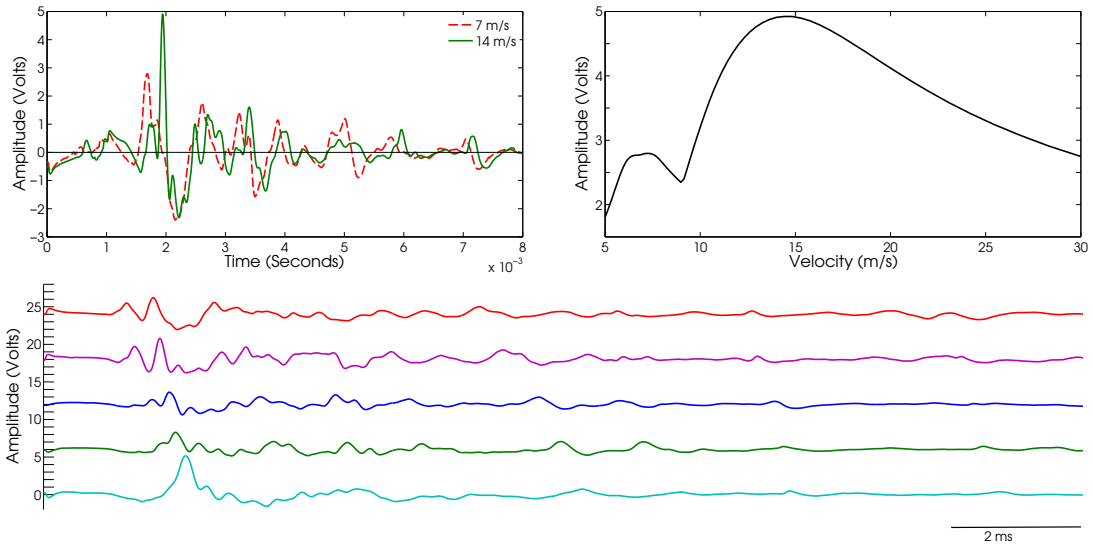


Figure 6.4.1: Electrical stimulation of the derma at 4 mA, the time domain recording **bottom inset** is synchronised with the stimulation pulse and offset by 1 ms. The dominant conduction velocities are at 7 m/s and 14 m/s as illustrated by the IVS **right inset**. The V_D waveforms corresponding to the peak velocities of 7 m/s and 14 m/s **left inset** show the location within the time domain of these dominant peaks. The ordinate scale in each case represents signal amplitudes after total system gain of 80 dB.

from 4.04 μV_{rms} to 7.31 μV_{rms} over the five channels. The observed *peak-to-peak* (pp) amplitudes for single APs were in the range 33.34 μV_{pp} to 65.49 μV_{pp} with larger amplitudes observed for the faster APs as is to be expected based on theoretical models [13]. SNR values (pp/rms) were therefore in the range 13.2 dB to 24.2 dB.

Figure 6.4.2 a) shows the VSD (or number of APs within each velocity band) for all of the recordings in the resting state. The VSD was computed using the mean level of activity as measured in each of the ten recordings. Error bars were fitted showing the standard deviation from the mean level of activity from all ten recordings. As already noted, the majority of neural activity falls in the range 5 m/s – 20 m/s and so the analysis has been restricted to this range (a velocity step of 1 m/s has been employed throughout). The VSD histogram clearly indicates a bimodal peak. The two dominant peaks are located at 10 m/s and 12 m/s respectively, with a consistent dip in activity at 11 m/s seen in each recording. The average number of APs identified during each 250 ms recording was 160 with a standard deviation of 7.95 APs.

6.4.3 Naturally Evoked (Physiological ENG): Cutaneous Skin Sensation

Recordings of cutaneous skin sensation were made while lightly stroking the L5 dermatome. As in the resting state, 10 recordings of duration 250 ms were made at 30-second intervals. The input-referred noise floor was measured once again during a 5 ms period of no observable neural activity and the measured values ranged from 4.00 μV_{rms} to 8.31 μV_{rms}. The

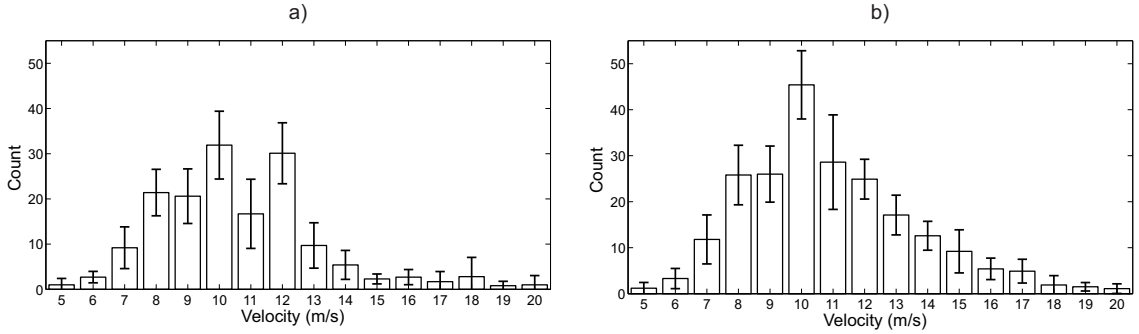


Figure 6.4.2: VSD histograms computed for each 2.5 second composite recording made without (a) and with (b) manual stimulation of the derma. The VSD is shown with mean activity levels for each velocity band, error bars show the standard deviation from all ten recordings.

peak-to-peak signal values for a single AP were in the range of 34.95 μV_{pp} to 86.49 μV_{pp} corresponding to SNR values of 12.5 dB to 26.7 dB. These values are very similar to the previous measurements.

Figure 6.4.2 b) shows the VSD histogram for the recordings made with cutaneous stimulation of the L5 dermatome. These are in the same format as Figure 6.4.2 a), for the resting case. Since the exact location of the stimulation event within the time record was unpredictable, each recording was inspected visually to ensure that one (and only one) event had been captured. The VSD histogram in Figure 6.4.2 b) clearly indicates that the distribution of axon firing has changed from a bimodal distribution to a single peak located at 10 m/s. The average number of APs identified during each 250 ms recording was 220.7 with a standard deviation of 27.07 APs (compared to 160 with a standard deviation of 7.95 APs for the resting state).

6.4.4 Validation

In order to validate the VSD process, the conduction velocities of 50 APs selected at random were measured in the time domain by hand, the average delay as measured from the peak of the AP from one channel to the next was used to compute the conduction velocity. The results were compared with the output from the VSD processor. In all cases there was exact agreement between the conduction velocities measured by the two methods.

6.4.5 Spike Density Functions

An alternative method of presenting the output of the VSD process is to plot the time record of the output of the VSD gating function (a spike train). This was done for both the resting and stimulated data and the results are shown in Figure 6.4.3. Since the outputs are velocity dependent, a velocity of 10 m/s was chosen since earlier experiments

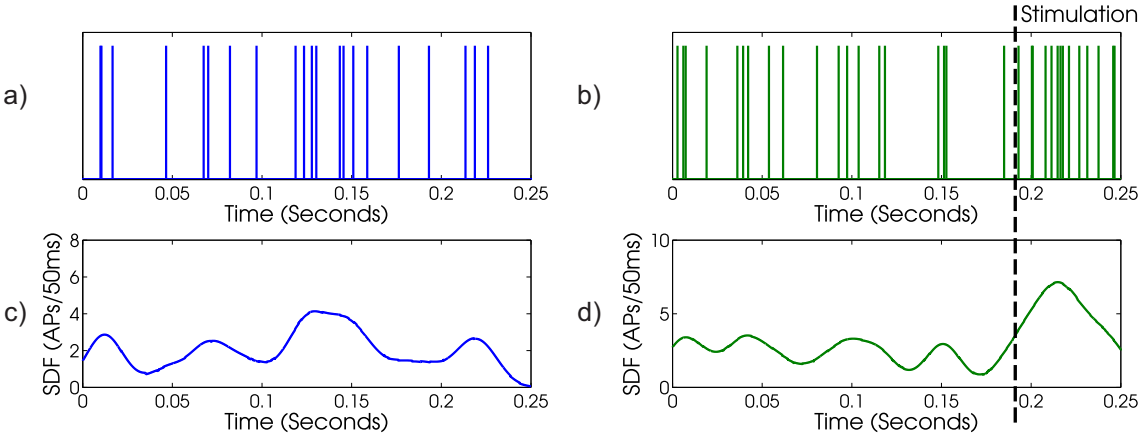


Figure 6.4.3: Spike trains (top) and spike density functions (bottom) for APs propagating at 10 m/s over a 250 ms recording with (right) and without (left) manual stimulation of the derma. The start of the stimulation event is indicated in the figure and corresponds to a visible increase in firing rates at a conduction velocity of 10 m/s.

demonstrated a clear increase in apparent firing rate during stimulation at that velocity. In order to represent the spike train as a continuous function, the spike train was convolved with a smooth and continuous kernel function [14].

Figures 6.4.3 a) and b) show the spike trains for APs propagating with a velocity of 10 m/s with and without cutaneous stimulation of the L5 dermatome respectively. The stimulation event occurred at approximately 200 ms from the start of the second recording. Figures 6.4.3 c) and d) show the spike trains that have been smoothed by convolution with a normalised Gaussian window 50 ms in length [14]. Both recordings show an oscillatory background firing rate, approximately limited to between 1 and 4 APs per 50 ms. At the onset of stimulation there is a significant jump in the firing rate to approximately 7 APs per 50 ms.

6.4.6 Results from Damaged Nerve

During initial experiments data were recorded that appeared to show almost instantaneous conduction of the AP along the length of the nerve, with signals appearing to be almost common mode in nature but still modulated by physiological factors such as stroking or electrical stimulation. It was suspected that the nerve had been damaged during the surgical procedure and was no longer propagating APs in the normal manner. In addition to the apparently instantaneous conduction of APs there was a successive reduction in AP amplitude in the orthodromic direction of propagation. It was hypothesised that the observed signals may have been the result of purely passive conduction (i.e without active regeneration of the APs, a process that in fact slows the conduction velocity as demonstrated in Chapter 2).

This hypothesis was validated by comparing measured AP amplitudes to those predicted by the theory assuming purely passive conduction. Starting with the concept of the

Table 6.4.2: Extracted AP amplitudes for a passively conducting neuron.

Channel	Peak Amplitude (mV)	Normalised Amplitude
1	2.152	1
2	1.745	0.81
3	1.457	0.67
4	1.174	0.55
5	1.087	0.505

passively conducting neuron as a simple lumped model transmission line, the voltage at the end of the neuron V_{out} as a function of the injected current I_{in} and the membrane properties (r_m and c_m, the membrane resistance and membrane capacitance respectively) is approximated by [15]:

$$V_{out} = I_{in} r_m \left(1 - e^{-\frac{t}{\tau}}\right) \tag{6.4.1}$$

Where:

$$\tau = r_m c_m \tag{6.4.2}$$

Now considering just the peak amplitudes over a distance x:

$$V(x) = V_0 \left(1 - e^{-\frac{x}{\lambda}}\right) \tag{6.4.3}$$

Using the peak amplitudes taken from the measured data and starting with the first recording channel (the decay was in the orthodromic direction of propagation) the amplitude decay along the nerve was extracted (Table 6.4.2).

Exponential curve fitting tools were used to extract the parameters for an exponential decay in amplitude with fit $R = 0.998$, the values found were:

$$V(x)_{fit} = 0.99 e^{-66.51x} \approx e^{-\frac{x}{0.015}} \tag{6.4.4}$$

Thus, for the values recorded:

$$\lambda = 0.006883 \tag{6.4.5}$$

This value of the space constant λ is within expected values for a large myelinated nerve (typical value is ≈ 0.004 [15]). If the nerve had been regenerating APs then this time value would have been much larger due to the time delay introduced by the regeneration process. The available evidence suggests that the recorded signals were a result only of passive electrical conduction along the nerve.

Fundamentally this evidence strongly suggests that the nerve had become damaged during the surgical procedure. In practice, it would be difficult to define the precise mechanism that damaged the nerve although in further experiments using more flexible hooks the effect was not observed. It is important to note that this failure mode would not be detectable using single electrode systems, only the use of a longitudinally spaced multiple electrode system permits this kind of analysis.

Table 6.5.1: The statistical significance as measured using the t-test of the variation between the individual experiments.

Velocity	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
P(%)	78	44	22	22	8.6	0.2	1.6	0.7	0.5	0.1	0.3	2.2	0.3	58	17	89
H	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0

6.5 Discussion

6.5.1 Validity of Results

Within the nervous system, information is encoded in terms of neuronal firing rates and so an increase in the amplitude of the stimulus results in a correlated increase in the rate of AP generation [16]. In order to evaluate the statistical significance of the variation between the individual experiments a paired t-test was performed using all available data for each velocity band [17]. Data from the resting and cutaneous stimulation experiments were compared and the results are given in Table 6.5.1 (the threshold for the null hypothesis was set to 5%). The variable **P** represents the probability of observing a test result as extreme as the observed value if the null hypothesis is true. **H** is a Boolean variable indicating rejection of the null hypothesis when **H** is set to 1. The null hypothesis was rejected for the velocities in the range 10 m/s − 17 m/s suggesting that the cutaneous stimulation produced an observable change within these velocity bands.

The velocity band 10 m/s−17 m/s is within the accepted range of conduction velocities for the $A\delta$ (5 m/s−30 m/s) afferent fibres in *rat*, which are responsible for light touch sensation [18]. Typically motor signals would be carried by efferent fibres, however because the signals recorded within this study were purely afferent in nature, the most likely candidate for the source of the increased activity is the $A\delta$ group of sensory nerve fibres.

During each recording there was some observed variability of both the signal and noise amplitudes from one electrode channel to another. The most likely explanation for this is the variation in extracellular resistance along the length of the fascicle that can affect the amplitude of the recorded signals [19]. When using cuffs or micro-channels the position of the nerve is tightly constrained and so the interelectrode impedances are fairly consistent along the array. This is not the case with simple wire hook electrodes where the nerve is only resting on each hook and so the longitudinal extracellular impedance was likely to have been variable along the array.

6.5.2 Interference and Noise

Tripolar amplifier configurations have been shown to reduce common mode interference signals - such as the *electromyogram* (EMG) - that are typical of *in vivo* recordings of this nature [20]. In this study the use of insulating paraffin oil around the fascicle and the application of a band-pass filter reduced interference from external sources to an acceptable

level so the use of a tripolar recording configuration was not warranted. In addition the animal was anaesthetised throughout and so there was little or no interface from EMG. In addition, the Humbug 50 Hz noise eliminators used in combination with the Faraday cage were effective in reducing interference from nearby electrical equipment and so no further processing was required. However, fundamentally the VSD could be extracted from recordings made with tripolar (double-differential), instead of single-differential signals.

At a first glance it appears that the VSD method increases the velocity selectivity, as each AP has been assigned a singular velocity as opposed to a detailed velocity spectrum, however this is not the case. The intrinsic selectivity has been shown to be a function primarily of the electrode geometry and the sample rate of the data acquisition system [20]. VSD effectively quantises the peak of each IVS to a discrete velocity and associates with each velocity an amplitude proportional to the number of occurrences of that velocity in the data set. In addition the width of the bins of the histogram is pre-set to the velocity resolution of the system, 1 m/s in this case. As an example, if VSD were applied to the CAPs recorded during electrical stimulation (as shown in Figure 6.4.1), only the conduction velocity with the largest amplitude, 14 m/s in this case, would be correctly identified. The results of electrical stimulation demonstrate the difficulty of isolating APs that are closely overlapping in time. VSD is better suited to sequences of nonoverlapping APs where there is only a single peak within the IVS.

This limitation suggests a constraint on the maximum number of APs that may be correctly detected within a given time window. Recent experiments in microchannel structures, for example, have observed approximately 100 myelinated axons in a channel of dimensions 100 μm^2 [6]. In this study individual axons were observed with relatively low firing rates and so the APs were well separated in time. However if all 100 axons were firing at 10 Hz, for example, then the composite observed signal would contain 1000 APs per second and it is very likely that overlapping APs would occur. In this study the duration of recorded APs was approximately 1 ms and so the theoretical, although statistically unlikely, limit for reliable detection is 1000 APs per second. In a practical application the size of fascicle that must be isolated in order to significantly reduce the likelihood of overlapping APs is related to two factors: the expected maximum firing rate of each axon and the number of axons within the fascicle. At present it is not known to what extent naturally occurring APs would overlap, or to what extent this would affect the methods proposed in this chapter.

6.5.3 Applicability to Chronic Recordings

Although hooks are inappropriate electrodes for chronic study, they were chosen for this acute experiment to test the new VSD method. The use of hook electrodes provides time domain recordings with high SNR allowing AP propagation velocity to be calculated

by hand and does not require custom-designed amplifiers. However, it should be noted that the signal processing techniques developed in this thesis can be readily extended to *multiple electrode cuffs* (MECs), provided that their lumens are small enough to give distinguishable spikes in the neurogram. Cuffs have a long proven history of stable chronic implantation in *man* [21]. Even more recent studies have shown that the use of micro-channel nerve interfaces that trap fine nerves or several separated fascicles should improve velocity selectivity, assuming more than three electrodes can be arranged in each channel [22].

The signal processing methods used to implement VSR and VSD are, as previously noted, fundamentally simple systems with the ability to be implemented in a low power real time configuration. This is in contrast to existing neural recording systems that generally employ statistical methods, such as *principal component analysis* (PCA) and clustering. These methods not only require intensive computation but also cannot generally be operated in real-time since they require training [4]. The requirement for prior knowledge in spike sorting systems varies from one method to another, but generally speaking a good deal of information about the shape of the various AP waveforms is required before online processing can occur [4]. By comparison the VSD system described in this thesis has a limited number of free variables, the width of the centroid filter (which is non-critical) and the noise floor of the recordings which can be estimated in real-time. VSR systems are therefore more suited to applications requiring implantation and real-time operation than approaches based on conventional pattern processing. A more detailed comparison of the implementation costs of the various methods is required before a definitive comparison can be made.

6.6 Conclusions

A method for extracting neuronal firing rates from physiological ENG based on conduction velocity has been demonstrated using *in vivo* recordings in *rat*. Simple wire hook electrodes were used to form a short recording array in which a micro-dissected but intact fascicle was placed. Data were recorded using commercially available amplifiers and data converters before being processed using basic operations in MATLAB. This method generates a detailed overview of the firing rates of neurons based on their conduction velocity and direction of propagation. Changes within the firing rates for particular velocities were observed during both electrical and mechanical stimulation of the L5 dermatome and recorded signal amplitudes were sufficient to negate the use of averaging or more complex recording arrangements. Although it was shown that this method is directly applicable to physiological ENG, it remains to be investigated whether it is transferable to chronically implanted electrode structures such as cuffs or micro-channels.

This is the first experiment in which velocity selective based methods have been applied successfully to the analysis of physiological ENG.

References

[1] M. Schuettler, V. Seetohul, N. J. M. Rijkhoff, F. V. Moeller, N. Donaldson, and J. Taylor, "Fibre-selective recording from peripheral nerves using a multiple-contact cuff: Report on pilot pig experiments.," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 3103–6, Aug. 2011.

[2] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[3] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. N. Donaldson, and J. T. Taylor, "An Enhancement to Velocity Selective Discrimination of Neural Recordings: Extraction of Neuronal Firing Rates," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2014, pp. 4111–4114, Aug. 2014.

[4] S. Gibson, J. W. Judy, and D. Markovic, "Spike Sorting: The First Step in Decoding the Brain," *IEEE Signal Processing Magazine*, vol. 29, pp. 124–143, Jan. 2012.

[5] K. Famm, B. Litt, K. J. Tracey, E. S. Boyden, and M. Slaoui, "Drug discovery: a jump-start for electroceuticals.," *Nature*, vol. 496, pp. 159–61, Apr. 2013.

[6] D. J. Chew, L. Zhu, E. Delivopoulos, I. R. Minev, K. M. Musick, C. A. Mosse, M. Craggs, N. Donaldson, S. P. Lacour, S. B. McMahon, and J. W. Fawcett, "A microchannel neuroprosthesis for bladder control after spinal cord injury in rat.," *Science translational medicine*, vol. 5, 2013.

[7] E. Delivopoulos, D. J. Chew, I. R. Minev, J. W. Fawcett, and S. P. Lacour, "Concurrent recordings of bladder afferents from multiple nerves using a microfabricated PDMS microchannel electrode array.," *Lab on a chip*, vol. 12, pp. 2540–51, July 2012.

[8] I. R. Minev, D. J. Chew, E. Delivopoulos, J. W. Fawcett, and S. P. Lacour, "High sensitivity recording of afferent nerve activity using ultra-compliant microchannel electrodes: an acute in vivo validation.," *Journal of neural engineering*, vol. 9, p. 026005, Apr. 2012.

[9] J. Taylor, N. Donaldson, and J. Winter, "Multiple-electrode nerve cuffs for low-velocity and velocity-selective neural recording," *Medical & Biological Engineering*, vol. 42, pp. 634–643, Sept. 2004.

[10] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "The theory of velocity selective neural recording: a study based on simulation.," *Medical & biological engineering & computing*, vol. 50, pp. 309–18, Mar. 2012.

[11] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "A summary of the theory of velocity selective neural recording.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2011, pp. 4649–52, Jan. 2011.

[12] J. S. Carp, A. M. Tennissen, and J. R. Wolpaw, "Conduction velocity is inversely related to action potential threshold in rat motoneuron axons.," *Experimental brain research*, vol. 150, pp. 497–505, June 2003.

[13] K. G. Pearson, R. B. Stein, and S. K. Malhotra, "Properties of action potentials from insect motor nerve fibres.," *The Journal of experimental biology*, vol. 53, pp. 299–316, Oct. 1970.

[14] A. Szcs, "Applications of the spike density function in analysis of neuronal firing patterns," *Journal of neuroscience methods*, vol. 81, pp. 159–167, 1998.

[15] R. Carpenter, *Neurophysiology.* Arnold Publishers, fourth ed., 2003.

[16] H. Milner-Brown, R. Stein, and R. Yemm, "Changes in firing rate of human motor units during linearly changing voluntary contractions," *The Journal of physiology*, pp. 371–390, 1973.

[17] M. O'Mahony, *Sensory evaluation of food: statistical methods and procedures.* CRC Press, 1986.

[18] H. Gasser, "The classification of nerve fibers.," *Ohio Journal of Science*, vol. 41, no. 3, pp. 145–159, 1941.

[19] M. Rahal, J. Winter, J. Taylor, and N. Donaldson, "An improved configuration for the reduction of EMG in electrode cuff recordings: a theoretical approach.," *IEEE transactions on bio-medical engineering*, vol. 47, pp. 1281–4, Sept. 2000.

[20] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

[21] K. H. Polasek, H. A. Hoyen, M. W. Keith, R. F. Kirsch, and D. J. Tyler, "Stimulation stability and selectivity of chronically implanted multicontact nerve cuff electrodes in the human upper extremity.," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 17, pp. 428–37, Oct. 2009.

[22] J. J. FitzGerald, N. Lago, S. Benmerah, J. Serra, C. P. Watling, R. E. Cameron, E. Tarte, S. P. Lacour, S. B. McMahon, and J. W. Fawcett, "A regenerative microchannel neural interface for recording from and stimulating peripheral axons in vivo.," *Journal of neural engineering*, vol. 9, Feb. 2012.

Chapter 7

Pilot Experiments in *Pig*

7.1 Motivations and Overview

It was discussed in Chapter 3 that in spite of the many potential applications of *velocity selective recording* (VSR) it has until recently only been demonstrated with *electrically evoked electroneurogram* (ENG), i.e. *compound action potentials* (CAPs) in *pig* and *frog* [1, 2]. In Chapter 6 recordings were made of physiological ENG in *rat* using hook electrodes, and for the first time the methods of VSR (modified using the method of *velocity spectral density* (VSD)) were successfully applied to extract functional information from the recordings. These recordings demonstrated not only that it was possible to record *action potentials* (APs) from an array of hooks but also that it was possible to extract firing rates for individual neurons from the composite recordings using VSR and VSD.

This chapter presents preliminary results of the logical next step - i.e the first experimental validation of VSR using a chronically implantable electrode structure *in vivo*. Experiments were performed to obtain both electrically evoked and physiological ENG from the entire right *vagus* nerve of an adult female *pig*. The vagus nerves (left and right) are the tenth pair of cranial nerves that innervate the heart, lungs, upper digestive track and other organs of the chest and abdomen. The surgical aspect of this work was performed under the kind guidance of Dr. Thomas Nielsen at Aalborg University. The hook electrodes employed in previous experiments in *rat* are only suitable for acute experimentation. Hooks were chosen for the experiments in *rat* because they provide a substantially higher *signal-to-noise ratio* (SNR) than chronically implantable structures such as cuff electrodes and they are easier to implant [3]. In this study *multiple electrode cuffs* (MECs) were employed as both the stimulation and recording electrodes and both were attached to the entire (i.e un-dissected) right vagus nerve.

Unlike the previous recordings made in *rat* the expected amplitudes of the extracellular potentials (V_{ex}) recorded with an MEC are substantially lower than obtained from the same nerve attached to hook electrodes (typically, using cuffs, 1 μV - 10 μV for physiological ENG and 100 μV for CAPs). The previous experiments in *rat* demonstrated that VSR

and VSD are capable of extracting functional information in situations with individual channel SNR > 0 dB. In this chapter recordings of the physiological ENG with per channel SNR < 0 dB are reported and VSD is applied successfully to extract conduction velocity-based firing rates. The recording MEC contained ten electrodes that were connected in pairs to form nine bipolar recording channels before amplification, band-pass filtering and digitisation.

Two sets of experiments were performed, the first using a separate MEC to electrically simulate the nerve and the second to record purely physiological ENG. Results are presented from both cases and illustrate the successful stimulation and recording from the entire nerve using the two MECs. The experimental setup did not differ otherwise from one experiment to the other. Modulation of the neural signals was elicited by the natural variations in cardiac function and blood pressure and VSD was used successfully to extract individual APs from the physiological electroneurogram. During electrical stimulation the stimulation current was increased so as to recruit initially the faster axons and then the slower ones. For each experiment the data were recorded continuously over a two minute period. The data presented in this chapter have yet to be fully analysed and the results presented are from a preliminary analysis.

The validation of VSR using physiological recordings of the ENG from a chronically implantable electrode structure is a vital step towards creating a device suitable for advanced neuroprosthetics. Methods developed and validated using these tools pave the way for the design of real time neural signal processors that require minimal setup and utilise chronically implantable electrodes that have a long history of stable implantation.

7.2 Experimental Setup

7.2.1 Animal Preperation

All animal procedures were performed in accordance with the *Danish Animal Experiments Inspectorate* (no. 2013-15-2934-00753). In the experiment an adult female Danish Landrace *pig* (weight approximately 50 kg) was sedated with zoletil administered by the *intramuscular* route, transported to the hospital and then intubated and anaesthetized with sevoflurane and zoletil, also administered intramuscularly. An incision of approximately 20 cm was made on the right side of the trachea to expose the right cervical vagus nerve (the tenth cranial nerve), and a section of approximately 15 cm of the nerve was freed from surrounding tissues. The recording cuff electrode was placed on the nerve at the cranial end of the freed nerve section, while a stimulation cuff electrode was placed on the nerve at the caudal end of the freed nerve section.

The animal was grounded via the insertion of a 20 cm conductive rod under the derma of the abdomen that was connected to the recording amplifiers and the *electrocardiogram* (ECG) and the vascular blood pressure were recorded independently of the ENG for the

Table 7.2.1: Specifications of the high speed ADCs used throughout the experiments, the sampling rate was 100 kS/s.

Property	**Value**
Gain Drift	6 ppm/C
Offset Drift	29 μV/C
CMRR ($f = 60$ **Hz**)	100 dB
-3 **dB Bandwidth**	$>$ 500 kHz
Input Impedance	$>$ 1 GΩ
Noise	0.75 LSB
THD (20 $\mathbf{V}_{pp}$ **at** 10 **kHz**)	-85 dB
Crosstalk (20 $\mathbf{V}_{pp}$ **at** 1 **kHz**)	-100 dB

purposes of vital sign monitoring. After all surgical procedures had been completed, the anaesthesia was switched to a mixture of ketamine, midazolam, and fentanyl (administered *intravenously*) and the *pig* was left to stabilize for at least 20 minutes before performing the experiments. The animal was euthanised after completion of the experiments by a leathal dose of pentobarbitone.

7.2.2 Electrode and Amplifier Configuration

Both the stimulation and the recording cuff electrodes were produced according to the technique described by Haugland [4]. The stimulation cuff was approximately 12 mm long, with an inner diameter of 2.4 mm, and contained three 1 mm wide platinum-iridium ring electrodes with 5 mm centre-to-centre electrode spacing. The recording cuff was approximately 42 mm long, with an inner diameter of 2.4 mm, and contained eleven 1 mm wide ring electrodes with 3.5 mm centre-to-centre electrode spacing. The cuff electrodes were placed around the nerve by placing the cuff under the nerve, opening a slit on one side of the cuff, pushing the nerve through, and closing the cuff again. In order to minimize current leakage through the slit opening, a silicone sheet was placed around the cuff electrode with the opening in the reverse position, and the cuff was closed with ligatures at each end and at the centre. The recording cuff was connected to a custom made amplifier bank in a bipolar configuration and band-pass filtered from 10 Hz to 10 kHz, the overall amplifier gain was 80 dB.

7.2.3 Recording System

The amplified and filtered signal was passed to a set of high speed successive-approximation *analogue to digital converters* (ADCs) (NI9222 mounted in cDAQ-9178 by National Instruments, Austin, TX, USA) providing simultaneous sampling of all nine bipolar recordings with 16-bit resolution. The specifications of the converters are provided in Table 7.2.1.

The converters were connected to a battery operated portable laptop computer running LabView 2010 that logged each channel into a set of data files for offline analysis. No processing was performed on the data during the experiment. Each data file contained a time record and nine channels of raw data corresponding to each amplifier output. Online visual verification was provided by connecting a set of oscilloscope channels at the converter inputs, which provided a means to examine each recording in real time for conspicuous artefacts or noise.

7.3 Signal Processing

Offline data analysis was performed using MATLAB R2012b (The MathWorks, Natick, MA, USA). The raw data was imported into MATLAB using a custom file handler designed to extract automatically the individual channels from within the proprietary LabVIEW file format. Each of the nine bipolar signals formed from ten electrodes was sampled continuously for two seconds at a sample rate of 100 kS/s. The smallest possible delay was therefore 10 μs corresponding to a maximum velocity of 3.5 mm/10 μs = 3500 m/s and allows for a velocity resolution of 0.3 m/s at a velocity of 10 m/s. Each bipolar signal was time-shifted against the others by multiples of dt and then summed to form a single signal. Where the required delay was not a multiple of dt, each bipolar signal was time shifted by alternating values of dt such that the average delay over all nine channels was the desired value. The delay-and-add based *intrinsic velocity spectrum* (IVS) was used to extract the basic velocity information from the raw data, each recording lasting two minutes in this case [5, 6]. Further analysis was then performed using the VSD methods to create histograms and spike trains for individual velocity bands, the noise threshold used throughout the process was calculated by eye based on visual inspection of the recordings.

7.4 Data Analysis

7.4.1 Electrically Evoked Data

In order to validate the recording setup and to examine the velocity populations that were present within the vagus nerve, electrical stimulation was applied to the tripolar stimulation cuff. Over a two minute period a fixed width variable amplitude square pulse was repeatedly applied to the stimulation cuff with amplitude starting at 1 μA and increasing to 10 μA in ten steps. The stimulation current limits were chosen so that two distinct neural populations were selectively recruited, resulting in two observable CAPs. Figure 7.4.1 illustrates the time domain signals recorded from the nine bipolar channels of the recording MEC that was located approximately 5 cm from the stimulation cuff. The stimulation current for the CAPs of Figure 7.4.1 was 10 μA and the conduction velocities

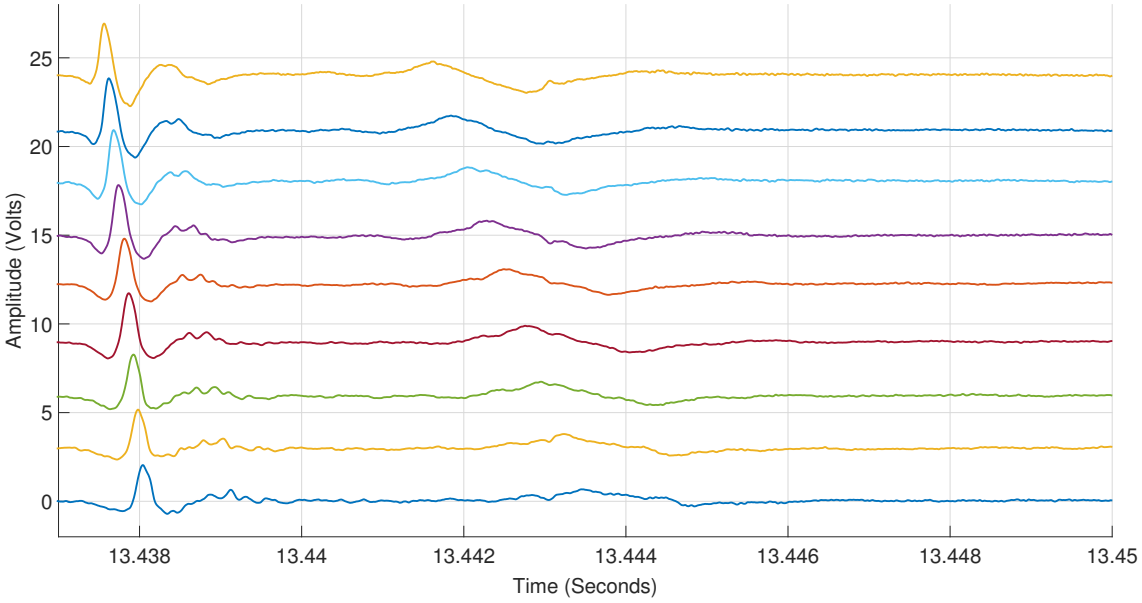


Figure 7.4.1: Time domain recording of the right vagus nerve starting from 1 ms after the onset of electrical stimulation that was supra-threshold for two main axonal diameters. Clearly visible are two CAPs propagating along the nerve with peak conduction velocities of 58 m/s and 15 m/s. The ordinate scale in each case represents the signal amplitude after application of 80 dB voltage gain and the offsets have been added artificially for presentational purposes.

for the two observable CAPs measured by hand is 58 m/s and 15 m/s with peak signal amplitudes of approximately 2 V and 0.5 V respectively. The offset in the data has been added artificially for the purposes of presentation.

The delay-and-add process (as detailed in Chapter 3) was applied to the electrically evoked data of Figure 7.4.1 and the IVS of Figure 7.4.2 was produced, the range of delay values applied corresponded to conduction velocities of 5 m/s - 85 m/s in steps of 1 m/s. The IVS of Figure 7.4.2 features two prominent velocity peaks located at 58 m/s and 15 m/s respectively and in accordance with the values measured by hand for the two CAPs of Figure 7.4.1. The velocity selectivities for the two electrically evoked CAPs are $Q_{v58} = 1.16$ and $Q_{v15} = 0.93$ respectively. In order to increase this selectivity artificially, band-pass filters were applied after the delay-and-add process (an explanation of the theory behind this process may be found in Chapter 3). Figure 7.4.3 illustrates the velocity spectra corresponding to three different band-pass filters with centre frequencies of 2 kHz, 4 kHz and 6 kHz respectively. The filters had relative bandwidths of 20% and were implemented as 8^{th} order Butterworth filters based on 4^{th} order lowpass prototypes.

The issue of spectral images was discussed in Chapter 3 and in Figure 7.4.3 the lower velocity population of 15 m/s has been masked by a spectral image that is located at a conduction velocity of approximately 10 m/s for a filter with centre frequency 2 kHz and moves to a velocity of approximately 22 m/s for a filter with centre frequency 6 kHz. Accordingly it is not possible to discern the 15 m/s peak associated with the lower amplitude CAP of Figure 7.4.1. The peak at 58 m/s shifts to 55 m/s with increasing filter centre

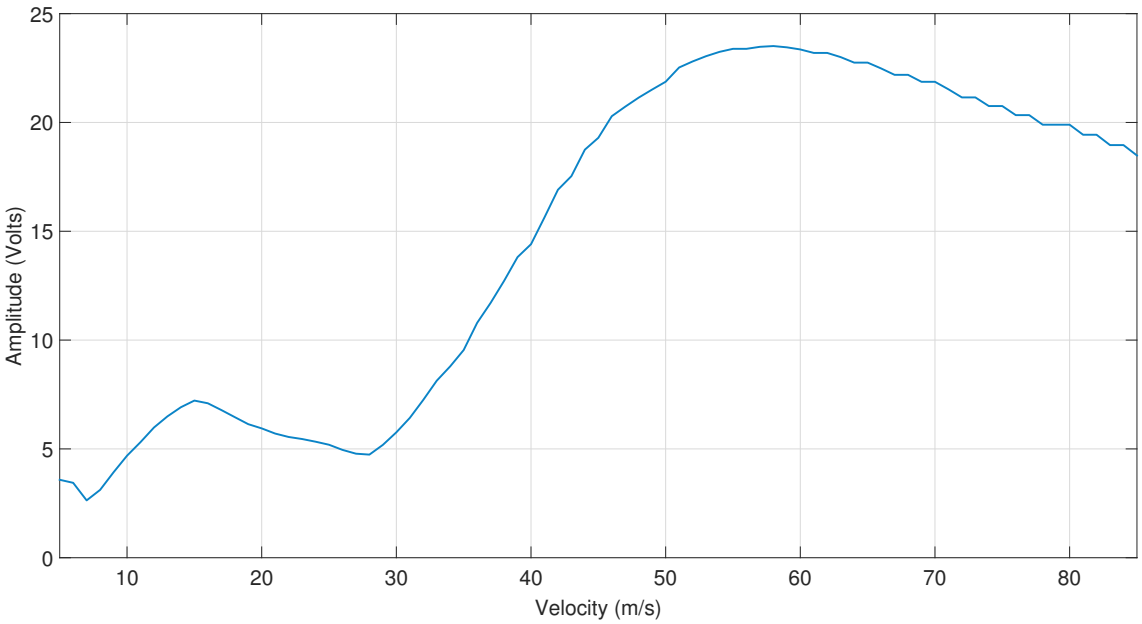


Figure 7.4.2: The IVS computed for the snippet of data shown in Figure 7.4.1 recorded from nine bipolar channels of data from the right vagus nerve 1 ms after the application of electrical stimulation from the stimulation cuff. The dominant conduction velocities are 58 m/s and 15 m/s respectively in accordance with the values measured by hand from the time domain recordings in Figure 7.4.1.

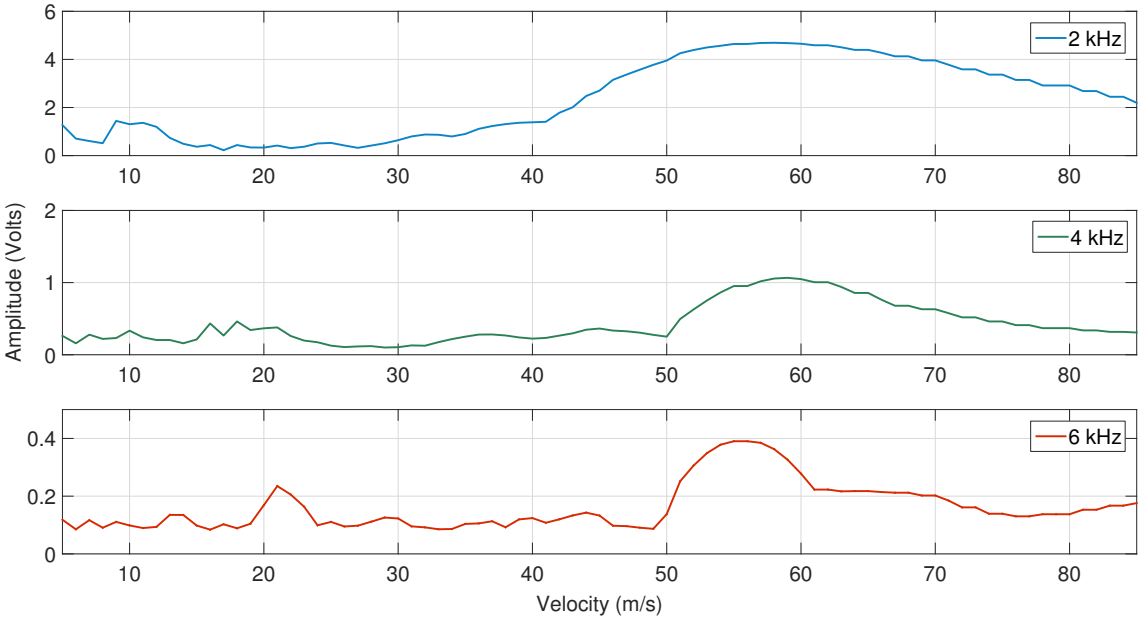


Figure 7.4.3: The band-pass filter enhanced IVS of the data recorded during electrical stimulation. From top to bottom the centre frequencies of the filters were 2 kHz, 4 kHz and 8 kHz respectively with relative bandwidths of 20%. The peak conduction velocity is 58 m/s, 59 m/s and 55 m/s respectively and the lower velocity component visible in the IVS of Figure 7.4.2 is masked entirely.

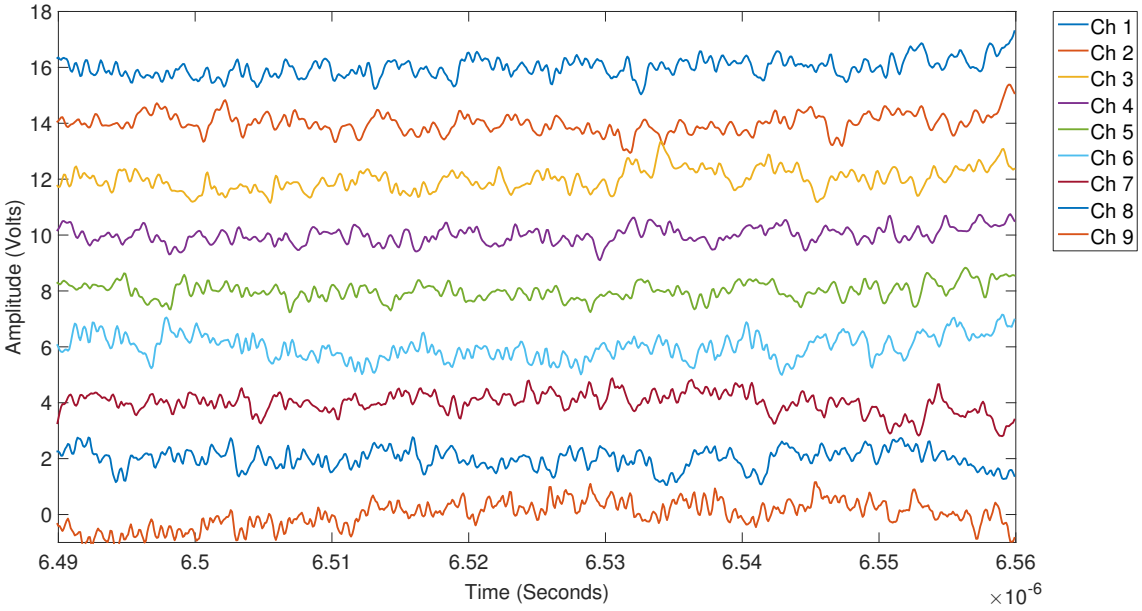


Figure 7.4.4: Time domain recording of physiological ENG from the right vagus nerve, there are no visually discernible features that could be attributable to ENG activity. The ordinate scale in each case represents the signal amplitude after application of 80 dB voltage gain, the offsets have been added artificially for presentation purposes. Channel one is proximal to the caudal end of the recording electrode.

frequency and the velocity selectivity for this peak increases from the intrinsic value of 1.16 to 5 for a centre frequency of 6 kHz. This represents an increase in velocity selectivity of approximately 4.31.

7.4.2 Physiological Data

Bipolar measurements of physiological ENG were made with the animal at rest - no stimulation was applied either electrically or manually. As with the electrical stimulation a continuous recording was made that lasted two minutes, the data were sampled concurrently over all nine bipolar channels at a sample rate of 100 kS/s with 16-bits of resolution. A short segment of each of the resulting nine channels of data is shown in Figure 7.4.4. Channel one is proximal to the *caudal* end of the recording electrode and channel nine proximal to the *cranial* end.

Examination of the traces in Figure 7.4.4 shows no visually discernible features that could be attributable to ENG activity. Results obtained with electrical stimulation of the vagus nerve distally from the recording electrode validated the recording system and so it was assumed that although individual APs were not visible it may be possible to use VSR to enhance the effective SNR. Delay-and-add was applied for velocities in the range 31 m/s to 45 m/s with a resolution of 2 m/s. These values were chosen after a more comprehensive sweep over a greater range of velocities revealed activity concentrated in this area. Figure 7.4.5 illustrates the post delay-and-add waveforms (V_D) for the same snippet of

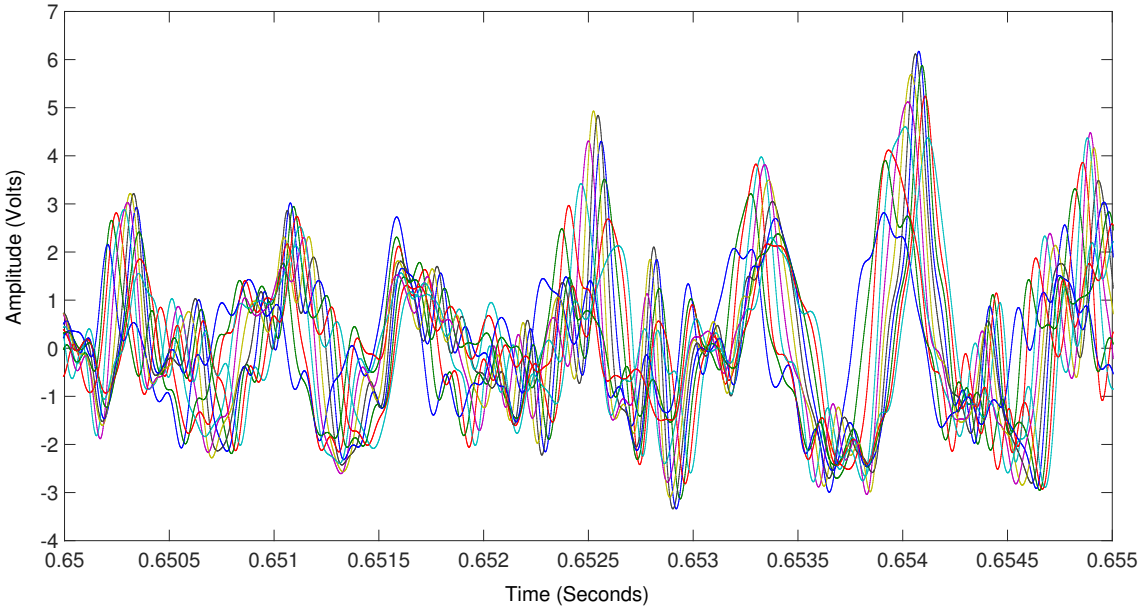


Figure 7.4.5: The post delay-and-add waveforms for the physiological recording of Figure 7.4.4 for matched velocities in the range 31 m/s - 45 m/s with an interval of 2 m/s, correlated peaks are clearly visible revealing several excited populations including one correlated peak with negative amplitude indicative of efferent neural activity.

physiological ENG shown in Figure 7.4.4. Clearly visible are correlated peaks that reveal several excited populations including one correlation with a negative peak indicative of *efferent* neural activity.

It was discussed in Chapter 3 that it is the interchannel correlation that is key to the delay-and-add process, whilst the time domain recording of the physiological ENG does not reveal any AP like features, the presence of features that correlate from one channel to another with a constant delay are strongly indicative of propagating waveforms. The VSD was computed for the snippet of raw data and is shown in Figure 7.4.6. VSD spike trains are shown for afferent activity within the velocity range 31 m/s – 45 m/s detected from the physiological ENG of Figure 7.4.4. The noise floor employed was set by eye to 2 V and there was a double response at 38 m/s near the end of the recording where two distinct peaks are visible in the delay-and-add waveforms of Figure 7.4.5. The ability to recover correlated data from nerve cuff recordings using an MEC results directly from the improvement in SNR provided by the delay-and-add process. This result is significant as it suggests that practical VSR systems can be used to record physiological data using MECs.

7.5 Conclusions and Summary

Both physiological and electrically evoked ENG have been successfully recorded *in vivo* from the right vagus nerve of a Danish landrace *pig*. Custom made MEC electrodes were

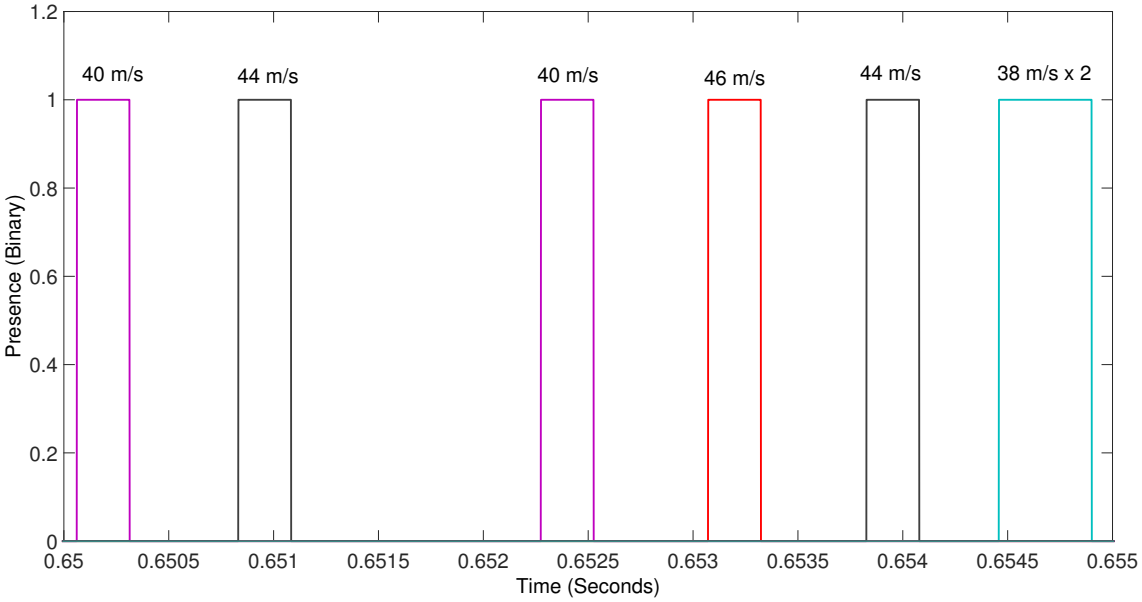


Figure 7.4.6: VSD spike trains for afferent activity within the velocity range 31 m/s - 45 m/s detected from the physiological ENG of Figure 7.4.4. The noise floor employed was set by eye to 2 V and there was a double response at 38 m/s near the end of the recording where two distinct peaks are visible in the delay-and-add waveforms of Figure 7.4.5.

used to form both stimulation and recording electrodes that were attached to the right vagus nerve. Data were recorded using custom built amplifiers and commercially available data converters before being processed using basic operations in MATLAB. Electrical stimulation was used to validate the recording setup and also to examine the range of conduction velocities that were support by the nerve, two dominant CAPs were observed with exceptional SNR for recordings made *in vivo* with cuff electrodes.

Recordings were then made of physiological ENG using the same recording setup. Although the individual channel responses showed no hint of the presence of APs, after the delay-and-add process it was possible to observe correlated peaks in the data that are indicative of APs propagating with biologically plausible conduction velocities in both the afferent and efferent directions. This is a crucial observation as it provides the first tentative proof of the practicality of the application of VSR to chronic recordings using structures such as MECs. The VSD method was applied to extract basic firing rates for the detected populations after the application of a simple band-pass filter, the noise floor chosen for the VSD was set by eye but could be automatically computed if required.

A full analysis of this data remains to be completed and the VSD method that has been demonstrated on a short time recording needs to be extended to extract the firing rates for the entire two minute recording of physiological ENG. This may then be compared to the recorded ECG and blood pressure waveforms and any possible correlations between the firing rates and the physiological signals detected.

References

[1] M. Schuettler, V. Seetohul, N. J. M. Rijkhoff, F. V. Moeller, N. Donaldson, and J. Taylor, "Fibre-selective recording from peripheral nerves using a multiple-contact cuff: Report on pilot pig experiments.," *Conference proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 3103–6, Aug. 2011.

[2] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[3] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. d. N. Donaldson, and J. T. Taylor, "A new method for spike extraction using velocity selective recording demonstrated with physiological ENG in Rat.," *Journal of neuroscience methods*, vol. 251, pp. 47–55, Aug. 2015.

[4] M. Haugland, "A flexible method for fabrication of nerve cuff electrodes," in *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 359–360, IEEE, 1996.

[5] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "The theory of velocity selective neural recording: a study based on simulation.," *Medical & biological engineering & computing*, vol. 50, pp. 309–18, Mar. 2012.

[6] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "A summary of the theory of velocity selective neural recording.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2011, pp. 4649–52, Jan. 2011.

Chapter 8

Design of a High Speed Multiple Electrode Cuff Simulation System

8.1 Motivations and Overview

Many of the recent developments in the field of *velocity selective recording* (VSR) have arisen from results obtained in pioneering experiments in animal models. *Ex vivo* experiments have been performed in *frog* [1, 2] and *in vivo* experiments presented in this work have been performed in both *rat* and *pig* [3]. Of course it is of great benefit to reduce the amount of time expended on animal experimentation, due to both the inherent costs and ethical constraints. It is the recommendation of the *United Kingdom Animals (Scientific Procedures) Act of 1986* that all experiments involving animals should follow a system of guidelines known as the "3 Rs", namely:

- **Replacement:** The substitution for conscious living higher animals of insentient material, where possible animal experiments should be avoided if technology allows.
- **Reduction:** To reduce the numbers of animals used to obtain information of a given amount and precision.
- **Refinement:** Refinement is any decrease in the incidence or severity of "inhumane" procedures applied to those animals that still have to be used.

Additional to these guidelines is the requirement to understand the costs and complexities involved with animal or human experiments. Not only are direct costs such as animal facilities and preparations expensive but there are also the indirect costs associated with planning and specialist equipment. The micro-surgical skills required for example to tease fine rootlets from a spinal cord into an electrode array require many years of training and experience. There is also a need for specialist low noise and high stability amplifiers, and high speed simultaneous data converters with equally good noise characteristics [3].

It is unlikely that there will ever be a complete substitute for biological experiments; there is simply too much about the nervous system and the interaction between tissue and electrodes that is not understood. It is possible however to use artificially generated signals in order to evaluate the core functionality of a neural recording system, and such systems have already been developed that produce simulated *action potentials* (APs) that are representative of those recorded from a *multiple electrode cuff* (MEC) [4]. This kind of signal generator may be directly interfaced to the analogue front end of a prototype VSR system, so that the entire signal processing chain, from amplifier to discriminator, may be tested by generating artificial APs with different conduction velocities and levels of noise. The currently available system has proven to be beneficial in the design of new VSR processers (in particular it was used for the evaluation of the *time delay neural network* (TDNN) method introduced in Chapter 3 [5]) but data recently recorded from *rat* has highlighted a number of drawbacks that need to be addressed if such a simulator can be used for more extensive testing of a VSR system. The key issues are:

- **Interchannel Delay:** With existing signal generators it is only possible to change the interchannel delay (and thus the apparent conduction velocity) over all the channels at once. This does now allow for the exploration of the effects of varying interchannel delays, such as those observed in biological recordings, on the effectiveness of the recording system.
- **Simulated Potentials:** The APs produced by existing signal generators are based on a simple model of the AP that does not take into account factors such as the tripole transfer function or the electrode impedance, and so the extent to which they represent physiological APs is open to debate. Additionally the sample rate of the system proposed by Al-Shueli et al. is 40 kS/s, whilst adequate to meet the Nyquist criterion for the widely accepted bandwidth of APs (approximately DC$-$10 kHz [6]) is too low to produce detailed reproductions and limits the available velocity selectivity.
- **Output Impedance:** When testing analogue amplifiers it is important to be able to experiment with different impedances, as different electrode configurations and animal preparations have significantly different electrical conductivity. Existing hardware systems do not allow for any variation in the output impedance. There are however more complex models that attempt to reproduce accurately the axon to electrode interface using an artificial axon [7]. Such systems provide a high level of realism and accuracy but at an increased cost and may also be less configurable.

In this chapter a new signal generator is described that overcomes these key limitations of existing systems. The new system has been designed and implemented on a Cyclone II EP2C35F672C6N *field programmable gate array* (FPGA) that is programmed and controlled via a dedicated onboard *universal serial bus* (USB) interface to a desktop or portable computer. Unlike existing systems, the AP waveforms are stored within the

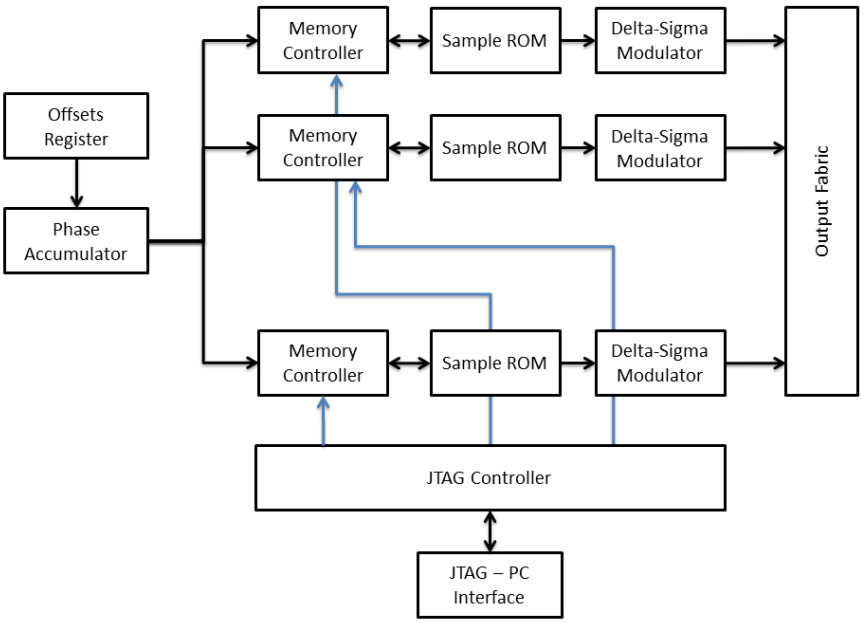


Figure 8.2.1: Simplified block diagram of the artificial AP generator, the blue interconnects represent the JTAG programming interface that is used to dynamically confgure the contents of the sample memories via the memory controllers.

FPGA so that once configured the generator may be disconnected from the desktop computer and can run as a standalone module. A set of scripts read data from a memory file stored on the computer and transmit this data into the FPGA, once configured the FPGA plays back this data at a set of pre-configured outputs, a simple set of filters and digital potentiometers are utilised to provide a high quality analogue representation of the APs with a customisable and realistic output impedance.

8.2 Implementation

The principle behind the proposed signal generator is similar to that of most arbitrary function generators; there is a bank of sample memories, one per output channel, and a *phase accumulator* that is incremented at the sample rate. The value of the accumulator is used as a pointer to the first sample memory. Each successive sample memory uses the accumulator value plus an offset for the address pointer, by varying the offset applied to each pointer the relative phase between each output channel can be individually controlled. Additionally there is a programming interface that is only connected when programming the sample memories, connecting the programming interface temporarily disconnects the sample memories from the rest of the system. Figure 8.2.1 provides an overview of the basic layout of the signal generator, the programming interfaces (which are implemented using a *joint test action group* (JTAG) interface and implemented using the onboard USB device) are shown in blue.

The sample values are retrieved from each sample memory synchronously to the master sample clock (the phase accumulator) and are fed into a single bit Delta-Sigma ($\Delta\Sigma$) modulator. The modulator reduces the multi-bit sample into a stream of digital pulses that can then be output using a dual-state pin on the FPGA. The motivation for using a Delta-Sigma modulator over more traditional methods such as *pulse width modulation* (PWM) will be explained in more detail.

8.2.1 Recordings Used for Playback

Existing AP signal generators produce AP waveforms that are based on theoretical models, such as the *transmembrane action potential* (TMAP) models that have been considered in Chapters 3 and 5 [4]. These relatively simple models may be superseded using the experimental recordings that have been made in *rat* (see Chapter 6). These recordings were made with a sufficiently high sampling rate (500 kS/s) and *signal-to-noise ratio* (SNR) ($>$ 13 dB) that they may be used as the basis for artificial waveforms. The use of simulated APs is driven by the need to replace animal experiments with a technological model and so ultimately the simulated APs must mimic the biological analogue as closely as possible. This is best achieved by using waveforms recorded from a biological source; the signals recorded from *rat* are strongly representative to those that would be present in an implantable system using MECs.

The bipolar data recorded from *rat* were examined by eye for well-defined APs with low levels of noise and interference. An example of such an AP is shown in Figure 8.2.2, once an AP was extracted from the raw recordings a set of MATLAB scripts were written that converted the raw data into a format suitable for downloading into the FPGA. Multiple APs were extracted from the recordings and allowed for up to five channels of realistic data (using the five separate channels recorded from *rat*). The discrete time samples recorded from *rat* are stored in a MATLAB array as double precision floating point numbers (according to the IEEE 754 format). The sample memories within the FPGA are designed as 16-bit fixed point values, so as to avoid the computational overhead associated with floating point arithmetic. A MATLAB script is used to convert the floating point samples into the Qs:0:15 fixed point format (signed, 0 decimal bits, 15 fractional bits) giving an effective precision of 0.00003. The values are then converted to hexadecimal before being sequentially written to a memory file that may be downloaded onto the FPGA.

8.2.2 JTAG Memory Interface

In order to provide read and write access to the sample memories on the FPGA an interface based around the JTAG standard was created. Altera Corporation, the manufacturer of the FPGA, provides a set of functions that may be used as *In System Sources and Probes*, these provide a set of signals within the FPGA that can be connected via the USB interface to a JTAG server running on a desktop or portable computer. The In System Sources and

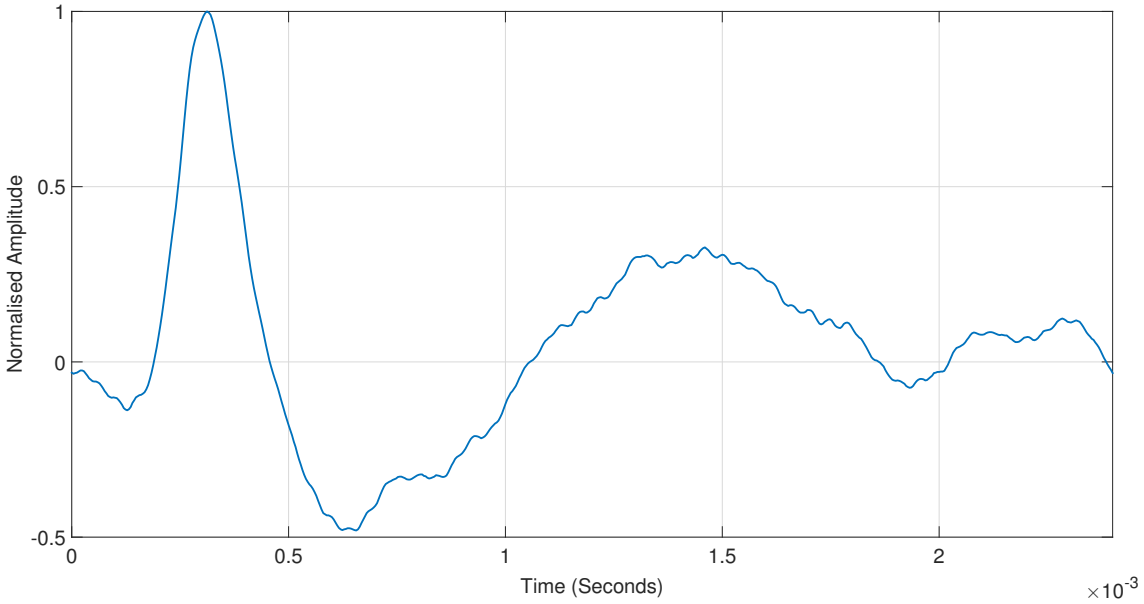


Figure 8.2.2: Single channel representative of APs extracted from bipolar recordings made in *rat* that was used as the template for the signal generator.

Probes editor within the Quartus software environment provides a means by which any signal internal to the FPGA may be controlled in real time.

A set of these sources and probes were used to provide an interface to the sample memories on the FPGA as well as to the offset register that controls the phase relationship between each channel, and thus the effective interchannel delay. The sample memories are accessed by both the Delta-Sigma modulators and the JTAG interface and so a multiplexor is used to select which interface is connected to the memories at a given time. The memories are programmed sequentially, with the memory selector bus defining which memory is currently connected to the JTAG interface. When the JTAG interface is disconnected the selector bus is cleared and each sample memory is connected by default to the phase accumulator and the Delta-Sigma modulators.

A typical programming sequence is as follows:

1. Set the memory selector bus to address of memory N.
2. Set memory select bit, indicating a JTAG read/write operations and switching the memory interface multiplexor into the JTAG mode.
3. Write data into the memory synchronously to the JTAG clock using the address, data and write enable registers.
4. Clear the memory select bit, returning control to the hardware.
5. Set the memory selector bus to the address of memory $N + 1$.

There is no constraint for the sample memories be programmed sequentially; they may be arbitrarily programmed with unique AP waveforms in any order. When each memory

is connected to the JTAG interface the inputs to the respective Delta-Sigma modulator are held at a digital zero, in this way the phase accumulator continues to run and so only the channel currently being programmed is effected by the memory operation in progress.

All of the sources and probes are operated coherently from the FPGA system clock in order to avoid timing violations when writing to the sample memories. The phase accumulator is operated from a user definable clock that is set to the system sample rate, when used with data recorded from *rat* this clock was set to a frequency of 500 kHz. The generator has been designed to be highly flexible and configurable. For example the sample rate, number of channels, memory width and memory depth are all user definable.

In order to automate the programming of the sample memories with the AP waveforms a set of scripts written in the Tcl scripting language are used. These scripts interface with the Quartus software environment and are responsible for abstracting the process of memory programming, for example functions are available within the scripts for downloading files that contain AP waveforms to the sample memories or for reading and writing single values from a particular memory. There are also run time functions that allow the user to change the interchannel delay on demand and in real time by transmitting values into the offset registers of the device. A suitable overview of the Tcl scripting language is available in [8]. The Tcl scripts make use of the Quartus In System Sources and Probes interface and require that the Quartus software is installed on the system, once configured however the generator may be unplugged from the system and left to operate in a standalone mode.

8.2.3 Electrode Spacing and Output Impedance

It was discussed in Chapter 6 that it is often difficult to obtain uniform interelectrode spacing when using hook type electrodes, or any other type of flexible electrode structure. The effect of the nonuniform spacing is to reduce the velocity selectivity of the system by introducing a variable interchannel delay and thus a spreading in the apparent conduction velocities of the APs. The signal generator described in this work provides the ability to change the delay between each of the output channels independently, this feature allows the effect of nonuniform electrode spacing to be examined in real time. The channel delays are stored within an offset register in the FPGA and the sample read from each sample memory is addressed as the sum of the phase accumulator and the individual channel offset. The resolution of the channel offset is one sample, and so for data recorded at 500 kHz the smallest possible interchannel delay is 2 μs. Another constraint derived from the recording structure is the effective electrode output impedance, this can vary considerably depending on the recording setup and if not properly matched to the input impedance of the amplifiers may cause instability and oscillations [9, 10]. In order to vary the output impedance of the signal generator each channel is connected to an AD5143

quad channel digital potentiometer (by Analog Devices, Massachusetts, USA) providing a possible range of output impedance from 60 Ω to 100 kΩ in steps of 390 Ω (256 different values).

8.2.4 Delta-Sigma Modulation

In order to convert the 16-bit sample values into a continuous time analogue waveform Delta-Sigma modulation is used. The motivation for using Delta-Sigma modulators to realise high-performance *digital to analogue converters* (DACs) is that it is often difficult to achieve a linearity better than an *effective number of bits* (ENOB) of about 14-bits for a DAC that is operated at the Nyquist rate [11]. The requirements for the analogue filtering of a Nyquist rate converter can become impractical as the anti-aliasing filter must have an upper cut-off frequency located at the Nyquist frequency with a maximally linear phase response in the pass band at this point. Such filters are problematic to realise efficiently and so are best avoided. Using Delta-Sigma modulation the constraints on the analogue filtering are relaxed and the filters are easily realisable. Using a fully digital Delta-Sigma modulator with an oversampled clock rate, such as that shown in Figure 8.2.3 allows a data stream with 16-bit word length to be converted into a single-bit pulse stream so that the baseband spectrum of the original signal is preserved. The noise shaping characteristics of the Delta-Sigma modulator shift the large amount of truncation noise that is generated in the loop to out of band frequencies. The single-bit pulse stream can then be converted by a two level DAC with (ideally perfect) linearity into an analogue signal. The truncation noise that is now out-of-band can be removed using a simple low pass analogue filter.

A two level DAC can be directly implemented as a digital output pin on the FPGA, with the two levels represented as logic 1 (3.3 V) and logic 0 (0 V). The up-sampling required to separate the spectral images of the baseband signal can be performed in MATLAB before the data is downloaded into the sample memories. While this is more memory intensive (as the stored data is no longer critically sampled) it relaxes the need to perform mathematically expensive interpolation within the FPGA. If however the number of channels, or the length of the recordings used for playback, is increased then the memory requirements of the system may be exceed; it was shown in Table 8.2.1 that the sample memory is the single largest component in the design. One way in which this memory requirement can be reduced is to perform the interpolation required for the Delta-Sigma modulator within the FPGA.

Table 8.2.1 provides the FPGA utilisation for a fully implemented design with five channels of sample memory; while the size of the design is not particularly constrained it is clear that the largest single requirement is the sample memory. It is possible to extend the design to produce an arbitrary number of output channels, each with a customisable interchannel delay and output impedance.

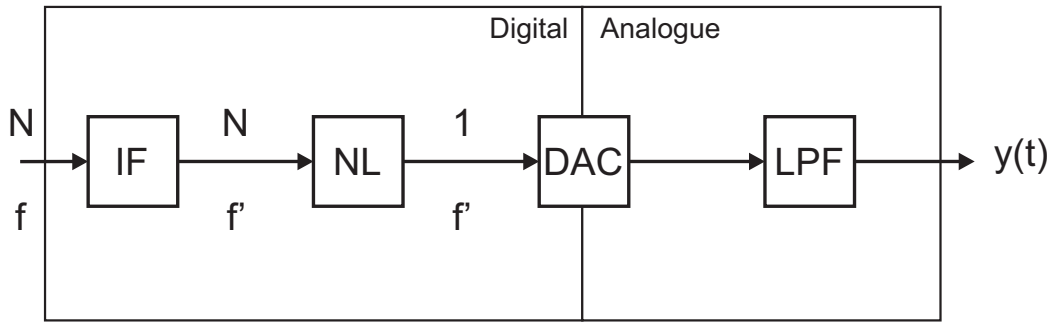


Figure 8.2.3: Block diagram of a Delta-Sigma DAC illustrating the **IF:** Interpolating Filter, **NL:** Noise shaping Loop, **DAC:** Digital to Analogue Converter & **LPF:** Low Pass Filter. The sample rates are f and f' and the bit width is reduced from N to 1.

Table 8.2.1: FPGA device utilisation for a five channel high speed simulation system and signal generator.

Component	Device Utilization
Total Logic Elements	890 / 33,216 (3%)
Total Combinatorial Functions	849 / 33, 216 (3%)
Dedicated Logic Registers	500 / 33,216 (2%)
Total Pins	50 / 475 (11%)
Total Memory Bits	51,200 / 483,850 (11%)

8.2.5 Polyphase Filtering and Cubic Interpolation Techniques

Sample Rate Conversion

In order to reduce the memory requirements for larger simulations, a separate interpolation model was designed that performs up-sampling and filtering using a polyphase filter structure and cubic interpolation. Cubic interpolation is a commonly used signal processing technique for increasing the sample rate of a signal, the interpolator has been constructed using a polyphase filter structure that provides many benefits over the traditional cubic interpolation structure and is decribed in detail. In many signal processing applications there is a need to perform sample rate conversion purely in the digital domain without an intermediate analogue filter. For a time series sampled at rate F, $x[n]$ there exists a linear but time varying system that will transform $x[n]$ to a new rate F'. Assuming a noninteger sample rate conversion, where the rate is decimated by a ratio M and interpolated by a ratio L then $y[m]$ (the new version of $x[n]$, sampled at F') is given by:

$$y[m] = \sum_{n=-\infty}^{\infty} g_m[n] x \left[\left\lfloor \frac{m}{L} \right\rfloor - n \right] \tag{8.2.1}$$

Where $g_m[n]$ represents the low pass filter required to remove spectral images that would occur due to aliasing.

In image processing applications, for example, this change in sample rate can be viewed as enlargement or reduction. Figure 8.2.4 illustrates this process for a decimator, $x[n]$

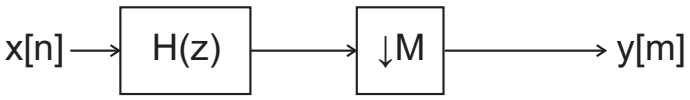


Figure 8.2.4: Block diagram for a down sampling converter, $x[n]$ represents the discrete time input, $H(z)$ the anti-aliasing filter and $y[m]$ the discrete time output at the new sample rate (a reduction of M).

represents the discrete time input, $g_m[n]$ the anti-aliasing filter and $y[m]$ the discrete time output at the new sample rate (a reduction of M).

Up-Sampling

In order to define the polyphaser filter structure required to perform interpolation it is necessary to begin by stating the general form for the time domain relationship for a $1 - to - L$ interpolator that can be derived from Equation 8.2.1:

$$y[m] = \sum_{-\infty}^{\infty} g_m[n] x \left[\left\lfloor \frac{m}{L} \right\rfloor - n \right] \tag{8.2.2}$$

Where the low pass filter $g_m[n]$ is given by:

$$g_m[n] = h[nL + m \oplus L] \tag{8.2.3}$$

To generate each output sample $y[m]$, $m = 0, 1, 2, ..., L-1$ a different set of coefficients $g_m[n]$ are used. After L samples are generated, the coefficient pattern repeats; thus $y[L]$ is generated using the same set of coefficients $g_0[n]$ as $y[0]$, $y[L+1]$ uses the same set of coefficients $g_1[n]$ as $y[1]$ and so on.

In general, for output samples $y[rL]$, $y[rL+1], ..., y[rL+L-1]$ the coefficients $g_m[n]$ are multiplied by samples $x[r-n]$. The partitioned subsets, $g_0[n]$,$g_1[n], ..., g_{L-1}[n]$, of $h[m]$ can be identified with L separate linear, time invariant filters that now operate at the low sampling rate F. For simplicity the time-invariant filters will be referred to as $p_0[n]$ *etc.* These filters $p_p[n]$ are referred to as *polyphase filters.*

Figure 8.2.5 illustrates the principle of the commutator model of polyphaser filtering, note that the filters operate on the original (lower) sample rate associated with $x[n]$. For each input sample x[n] each of the L branches of the polyphase network contributes one nonzero output that corresponds to one of the L outputs of the network. The polyphase implementation of this process has the advantage that the filtering is always performed at the lower sampling rate and so it eliminates the need for large and complex anti-aliasing filters. The low pass filter $g_m[n]$ is ideally a perfect low pass filter with an infinitely sharp transition from the pass band to the stop band and linear phase. Of course such a filter is unrealisable due to the infinitely long impulse response that would be required. There are a number of different filter functions that have been proposed for interpolation purposes and one of the most often used is the cubic convolution function. The cubic convolution

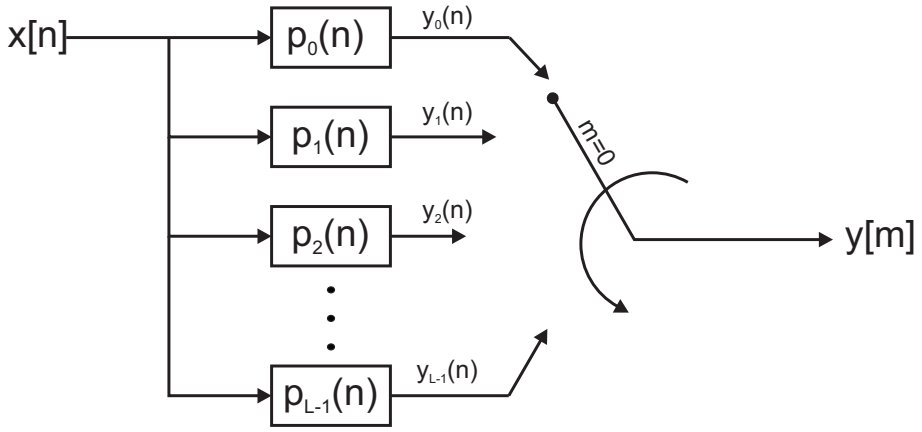


Figure 8.2.5: "Commutator Model" of a polyphase network interpolator.

is band limited and approximates the ideal low pass filter in the frequency domain (and thus a *sinc* function within the time domain). The kernel is given by [12]:

$$g_m[n] = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & 0 \leq |x| < 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & 1 \leq |x| < 2 \\ 0 & 2 \leq x \end{cases} \tag{8.2.4}$$

Where $n = x/x_s$ is the discrete time index of the continuous functions sampled at discrete points and a is a free parameter that controls the shape of the convolution kernel.

Keys has shown that the best approximation to the ideal low pass filter is given by $a = -0.5$ [12]. The time domain and frequency responses are shown in Figure 8.2.6, where the frequency response is given by:

$$H(\omega) = \frac{12}{\omega^2}\left(sinc^2\left(\frac{\omega}{2}\right) - sinc(\omega)\right) + a\frac{8}{\omega^2}(3sinc^2(\omega) - 2sinc(\omega) - sinc(2\omega)) \tag{8.2.5}$$

Where ω is the angular frequency.

The polyphaser filter structure defined in Figure 8.2.4 may now be used to perform the convolution of Equation 8.2.2 with the kernel function of Equation 8.2.4 and setting $a = -0.5$. Consider interpolation by a factor of four ($L = 4$), the values of the cubic convolution kernel at these points and hence the polyphase filter coefficients can be found using the following relationships [13]:

$$\alpha = \frac{n}{4} \tag{8.2.6}$$

$$n = 1, 2, 3 \tag{8.2.7}$$

$$P_0(0) = h(\alpha + 1) \tag{8.2.8}$$

$$P_0(1) = h(\alpha) \tag{8.2.9}$$

$$P_0(2) = h(1 - \alpha) \tag{8.2.10}$$

$$P_0(3) = h(2 - \alpha) \tag{8.2.11}$$

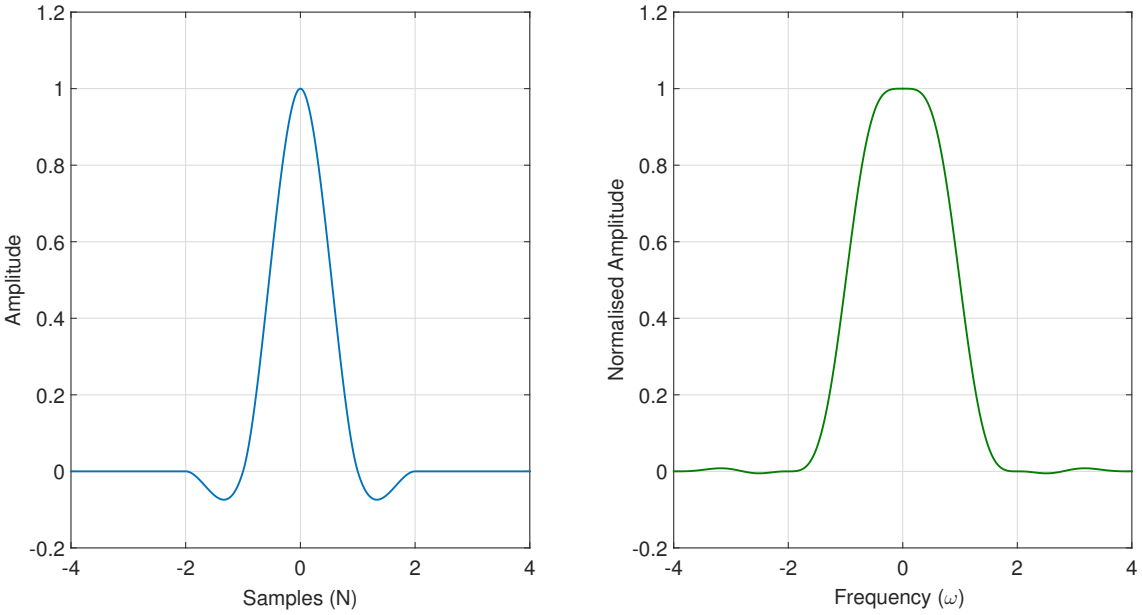


Figure 8.2.6: Time and frequency domain responses of the cubic convolution kernel for $a = -0.5$.

If the direct form implementation is used for each of the filters then it is clear from Figure 8.2.5 that the shift registers within each filter may be shared, saving overall memory consumption. As the interpolator only produces the sample points that are missing, an extra two taps need to be added to the commutator model that connect to the inner two taps of the input shift register respectively, to reproduce the original values at the output.

Down-Sampling

For a sense of completeness there exists a mirror structure for the process of decimation (down-sampling), it can be derived in a similar fashion as the interpolator structure by starting with the time domain relationship of the $M - to - 1$ decimator:

$$y[n] = \sum_{-\infty}^{\infty} h[k]x[nM - k] \tag{8.2.12}$$

The summation can be split into two by defining:

$$k = rM + \rho \tag{8.2.13}$$

And summing over the two variables $\rho = 0, 1, 2, ..., M - 1$ and $r = 0, \pm 1, \pm 2, ...$

$$y[n] = \sum_{\rho=0}^{M-1} \sum_{r=-\infty}^{\infty} h[rM + \rho]x[(n - r)M - \rho] \tag{8.2.14}$$

The input signals to the polyphase filters are:

$$x_p[n] = x[nM - \rho] \tag{8.2.15}$$

Where:

$$\rho = 0, 1, 2, ..., M - 1 \tag{8.2.16}$$

Using the relationship between the polyphase filter coefficients and the original filter coefficients:

$$P_\rho[n] = h[nM + \rho] \tag{8.2.17}$$

$$y[n] = \sum_{\rho=0}^{M-1} \sum_{-\infty}^{\infty} P_\rho[r] x_\rho[n - r] = \sum_{\rho=0}^{M-1} P_\rho[n] * x_\rho[n] \tag{8.2.18}$$

Where $*$ denotes convolution.

As in the case of the polyphase interpolator structure, a savings of factor M can be made in storage requirements for internal filter variables if the filters in the polyphase network are realised as transpose direct-form structures. Then the accumulator-shift register in each filter can be shared by summing the partial products of each respective tap of the polyphase filters with the respective taps of the other polyphase filters before storing them in the shift register. Aside from storage requirements, the number of multiplications per output sample is greatly reduced when using a polyphase network as the filters now operate on the lower sample rate so therefore the total computation rate in the system has been reduced by a factor of M or L.

8.3 Conclusions

An ongoing commitment to the "3 R" system (Reduction, Replacement & Refinement) as set out by the *United Kingdom Scientific Procedures act* requires that where possible simulation systems should be used to augment neural recordings made in animal models. The inclusion of a simulation system that is representative of a practical recording environment when designing neural signal processing systems, such as VSR, enables a reduction in the number of animal experiments required, a refinement in the results produced when animal experimentation is performed and the possibility to replace some experiments entirely. Currently available simulation systems are lacking in a number of key features that significantly reduce their effectiveness, such as the ability to simulate a variable electrode geometry, realistic electrode impedances and biologically plausible AP waveforms.

This chapter has described a new simulation platform that extends the existing technology using an FPGA to provide a high level of configurability and makes use recent recordings of physiological ENG in *rat* to address these key issues. It is unlikely that a simulation system will ever be capable of entirely replacing animal experimentation, especially in

such a cutting edge field as neural recordings. However, the use of systems such as the one described here helps to maximise the effectiveness of system such as VSR and thus minimise both the cost and the ethical burden associated with experimental neural recordings made in animal models.

References

[1] M. Schuettler, V. Seetohul, J. Taylor, and N. Donaldson, "Velocity-selective recording from frog nerve using a multi-contact cuff electrode.," *Conference Proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 2962–5, Jan. 2006.

[2] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[3] B. W. Metcalfe, D. J. Chew, C. T. Clarke, N. d. N. Donaldson, and J. T. Taylor, "A new method for spike extraction using velocity selective recording demonstrated with physiological ENG in Rat.," *Journal of neuroscience methods*, vol. 251, pp. 47–55, Aug. 2015.

[4] A. Al-Shueli, C. Clarke, and J. Taylor, "Simulated Nerve Signal Generation for Multi-electrode Cuff System Testing," *2012 International Conference on Biomedical Engineering and Biotechnology*, pp. 892–896, May 2012.

[5] A. Al-Shueli, C. Clarke, N. Donaldson, and J. Taylor, "Improved Signal Processing Methods for Velocity Selective Neural Recording Using Multi-Electrode Cuffs," *IEEE transactions on biomedical circuits and systems*, pp. 1–11, Sept. 2013.

[6] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

[7] O. Rossel, F. Soulier, S. Bernard, D. Guiraud, and G. Cathebras, "In-silico phantom axon: Emulation of an action potential propagating along artificial nerve fiber," in *VLSI (ISVLSI), 2015 IEEE Computer Society Annual Symposium on*, pp. 228–230, July 2015.

[8] P. Raines and J. Tranter, *Tcl/Tk in a Nutshell*. 1999.

[9] G. E. Loeb and R. A. Peck, "Cuff electrodes for chronic stimulation and recording of peripheral nerve activity.," *Journal of neuroscience methods*, vol. 64, pp. 95–103, Jan. 1996.

[10] M. Hansen, M. K. Haugland, and F. Sepulveda, "Feasibility of using peroneal nerve recordings for deriving stimulation timing in a foot drop correction system.," *Neuromodulation : journal of the International Neuromodulation Society*, vol. 6, pp. 68–77, Jan. 2003.

[11] R. Schreier and G. Temes, *Understanding Delta-Sigma Data Converters*. John Wiley & Sons, 2005.

[12] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Speech and Signal Processing*, no. I, pp. 1153–1160, 1981.

[13] L. R. Rabiner, *Multirate Digital Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1st ed., 1996.

Chapter 9

Validation of Velocity Selective Recording in the Context of Spike Sorting

9.1 Motivations and Overview

It was discussed in Chapter 3 that conventional analysis of extracellular recordings of the *electroneurogram* (ENG) employs one of several forms of *spike sorting*, methods based on neural data obtained from a single recording channel. Spike sorting methods are computationally intensive algorithms derived from conventional pattern processing techniques that are designed to classify *action potentials* (APs) morphologically [1]. Spike sorting methods attempt to correlate the underlying axonal diameter to the shape of individual extracellular APs. In addition to the computational expense, in virtually all cases training is required and so the spike sorting methods cannot generally be used in real time. These issues limit the usefulness of the methods especially their application to implanted devices such as the various types of "Bioelectronic Medicines" that are currently under discussion [2].

By contrast, *velocity selective recording* (VSR) and the related method of *velocity spectral density* (VSD) exploit the well-known relationship between axonal conduction velocity and diameter to discriminate between neural information from different sources. Unlike spike sorting, VSR requires several channels of data derived from the same neuron, a requirement that is conveniently fulfilled by the use of a structure such as a *multiple electrode cuff* (MEC), hook array or micro-channel interface structure. These devices are becoming more readily obtainable as standard components and are relatively easy to assemble surgically. Significantly, the real advantages of processing methods based on conduction velocity lie in the relatively simple and cheap signal processing needed and in the crucial fact that training is *not* required. The method is therefore well suited to applications requiring miniaturised low power implementations and real time operation.

It is worthwhile to consider and compare the two distinct approaches of spike sorting and VSR in order to understand the relative merits of each method. The purpose of this chapter is to present an outline comparison between these two methods using both simulated and experimental data obtained in a previous study from *frog* [3]. There are many ways in which the spike sorting process can be performed, and in order to undertake a comparison of the velocity based methods with spike sorting it is necessary to identify a typical spike sorting model. The key step in the spike sorting process is the feature extraction stage, whereby each AP, after identification and temporal alignment, is reduced to a set of descriptors that accurately describe its morphology. The *gold standard* for feature extraction is generally considered to be *principal component analysis* (PCA), a process by which each AP waveform is decomposed into a set of *principal components* - orthogonal basis vectors that allow the entire waveform to be characterised in a lower dimensional space.

In this chapter both conventional spike sorting methods and VSR are applied to simulated data and data recorded experimentally from an explanted nerve in *frog*. The results are compared and used as a validation that VSR and spike sorting, although based on very different methodologies, are able to classify extracellular ENG in the same way. Two different noise models are considered in the simulation study and the effect of the noise on the spike alignment process, and thus the classification process is examined. The effect of noise on the spike alignment process was described in detail in Chapter 4. Note that throughout this chapter and in keeping with existing literature the terms *AP* and *spike* are used interchangeably and refer to a single AP recorded extracellularly from an axon. The spike sorting process as a whole has been described in Chapter 3 and so only a brief overview will be given here.

9.2 Methods Employed for Spike Sorting

Spike sorting is a very general term that broadly speaking encompasses any method by which APs are classified based on their intrinsic morphology. Of course, it was shown in Chapter 2 that this morphology is directly related to the physical properties of the individual axons, a fundamental relationship that is the very basis for spike sorting. In order to examine the relationship between spike sorting and velocity based methods it is necessary to define a particular process for each, a process that is not only representative of the two methods but also enables a quantitative analysis. For the purposes of this study the following methods have been employed for spike sorting, each method has been chosen because it is considered to be the most commonly used during the sorting process. For spike extraction - the *nonlinear energy operator* (NEO), for spike alignment - the point of maximal slope, for feature reduction - PCA and for clustering - k-means clustering. These methods will be briefly examined in order to provide a qualitative comparison to VSR.

9.2.1 Nonlinear Energy Operator

The NEO is a method used to identify APs within time domain recordings, the NEO is designed to emphasise features within the time domain that have both a large instantaneous value and high frequency content. For a discrete time sampled signal the NEO takes the form of Equation 9.2.1.

$$\psi_d[x[n]] = x[n]^2 - x[n+1]x[n-1] \tag{9.2.1}$$

Where $x[n]$ is the discrete time signal that may be expressed as:

$$x[n] = Acos(\omega n + \theta) \tag{9.2.2}$$

Where A is amplitude, ω the angular frequency and θ the initial phase. Substitution of Equation 9.2.2 into Equation 9.2.1 gives:

$$\psi_d[x[n]] = A^2 sin^2(\omega n) \approx A^2\omega^2 \tag{9.2.3}$$

As the energy of an oscillation is proportional to the product of the square of the amplitude and the square of the frequency then Equation 9.2.3 represents the energy of $x[n]$ and so the NEO is often termed the *energy operator*. Examination of Equation 9.2.3 also shows that the NEO is suitable for the detection of a discontinuity or transient (such as an AP) because the output of the NEO is sensitive to instantaneous amplitude and frequency. Once the raw time domain signal has been processed using the NEO a simple linear threshold is applied, at each point that the NEO crosses this predetermined threshold a window of data containing the AP is extracted from the time domain waveforms and stored for further processing.

9.2.2 Principal Component Analysis

After the APs have been identified using the NEO they must be reduced to a set of features using PCA. PCA aims to reveal the variations in spike shape about the mean that have the largest variance. To an extent the variations of spike shapes among different axons are larger than variations of spike shapes from a single axon, therefore that PCA can help associate spike with individual neurons. PCA is a statistical procedure in which the recording of a spike waveform is considered as a vector in m-dimensional space, where m is the number of measurements (or samples) within the spike waveform. If there are n spike waveforms, each of length m samples, then a matrix of size n x m is formed where each of the n rows represents a spike that can be regarded as an m dimensional vector, a row vector in R_m. In R_m, PCA attempts to locate the best fitting linear combination of sets of orthogonal axes to replace the initial set of m axes in R_m.

The idea behind this process is to find a set of m' principal axes where $m' < m$ allowing the spikes to be adequately characterized on a smaller (m' - dimensional) space, while the

$m-m'$ dimensions may be ignored as describing noise. In PCA, the projections of the spike waveforms on the axis sought for needs to be as elongated as possible. That is to say that the variance of the projections needs to be as great as possible. The eigenvectors associated with the m' largest eigenvalues yield the best-fitting m' - dimensional subspace of R_m. This procedure essentially reduces each spike waveform to a set of *principal component scores* (PC Scores), or the weight associated with the spike and each *principal component*. Clustering of the first m' scores can then be used to classify each spike waveform. For a more detailed review on PCA and its applications see [4], a good review on the use of PCA for spike sorting is available in [5]. In this study the extraction of the PC scores was performed using the built in decomposition routines of MATLAB.

9.2.3 *k*-Means Clustering

Once the principal components for each AP have been computed it is possible to produce distinct classifications by observing clusters within the first N PC scores, where N is typically 2 - 3 [5]. In this study the clusters were automatically identified using the k-means clustering algorithm - a method of vector quantisation derived from signal processing in which n observations (in this case PC scores) are partitioned into k clusters. The process of identifying the clusters and partitioning each observation is iterative and is computationally difficult, formally it is NP-Hard.

Given a set of PC scores $(x_1, x_2, \ldots, x_N)$, where each score is a d-dimensional vector, k-means clustering aims to partition the N scores into k sets $S = S_1, S_2, \ldots, S_k$ where $k \leq N$, so as to minimise the within-cluster sum of squares such that:

$$\underset{s}{\text{minimize}} \sum_{i=1}^{k} \sum_{x \in s_i} ||x - \mu_i||^2 \tag{9.2.4}$$

Where μ_i is the mean of the points in S_i.

There are a number of ways of minimising the within-cluster sum of squares and in this study the built in function available in MATLAB was used. This function performs k-means clustering using Lloyd's algorithm (also known as *Voronoi iteration*) in which the PC scores are divided into partitions of well-shaped and uniformly sized convex cells. Voronoi iteration repeatedly finds the centroid of each set in the partition and then re-partitions the input according to which of these centroids is closest [6]. One of the drawbacks with the k-means clustering algorithm is that the value of k, the number of clusters, is required *before* clustering takes place. In the context of neural spike sorting this implies that the number of expected neuron classes be known in advance. Secondary to this is the exponential increase in complexity that arises when more than two dimensions are considered, although it has been shown that for neural spike sorting there is little benefit to considering more than the first three principal components (and thus dimensions required for clustering) [5].

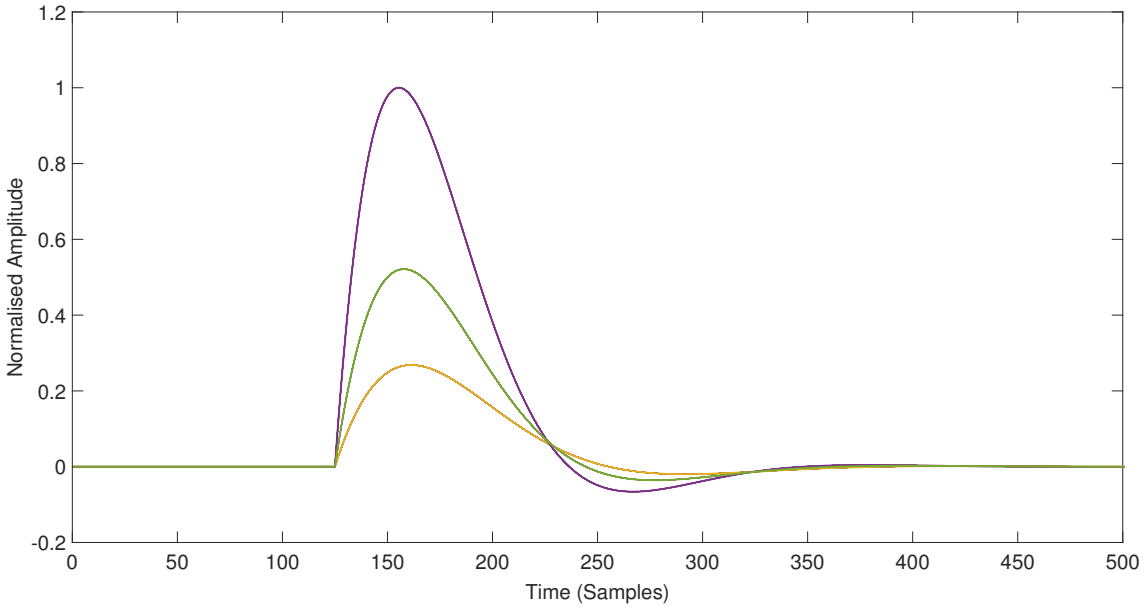


Figure 9.3.1: Overlay of the three AP models used for the simulation study, the time domain recording contained 160 APs selected at random from these models. The models correspond to three different axonal diameters (10, 15 & 22 μm) at 37 degrees Celsius.

9.3 Results for Simulated Data

9.3.1 Classification with Minimal Noise

In order to examine the basic operation of spike sorting using PCA as the feature extraction process a software simulation was performed using MATLAB. In this simulation APs were generated using the *transmembrane action potential* (TMAP) expression given in Equation 4.2.1 corresponding to recordings made with a bipolar amplifier configuration and ten electrodes, yielding nine channels of data. The effective interelectrode spacing was 7.5 mm. The simulated recording was one second in length and contained 160 APs selected at random from three different models corresponding to axonal diameters of 10, 15 and 22 μm. The *signal-to-noise ratio* (SNR) was modulated by the application of *additive white Gaussian noise* (AWGN) to yield an overall SNR of approximately 30 dB. The initial SNR was chosen so that there would be a visible variation within the classified clusters but also so that classification was trivial. Figure 9.3.1 illustrates the three AP waveforms used in this simulation, the relative position of each AP within in the time record was chosen to ensure that there were no overlapping APs. The effect of the axonal diameter on the peak extracellular voltage (Figure 9.3.1) is clearly visible with the largest amplitude recorded for the 22 μm axon (with a corresponding conduction velocity of approximately of 41 m/s).

The entire one second recording was analysed using standard delay-and-add in order to extract the *intrinsic velocity spectrum* (IVS) for a range of velocities from 5 m/s to 60 m/s with a resolution of 1 m/s. The IVS shown in Figure 9.3.2 clearly shows three distinct

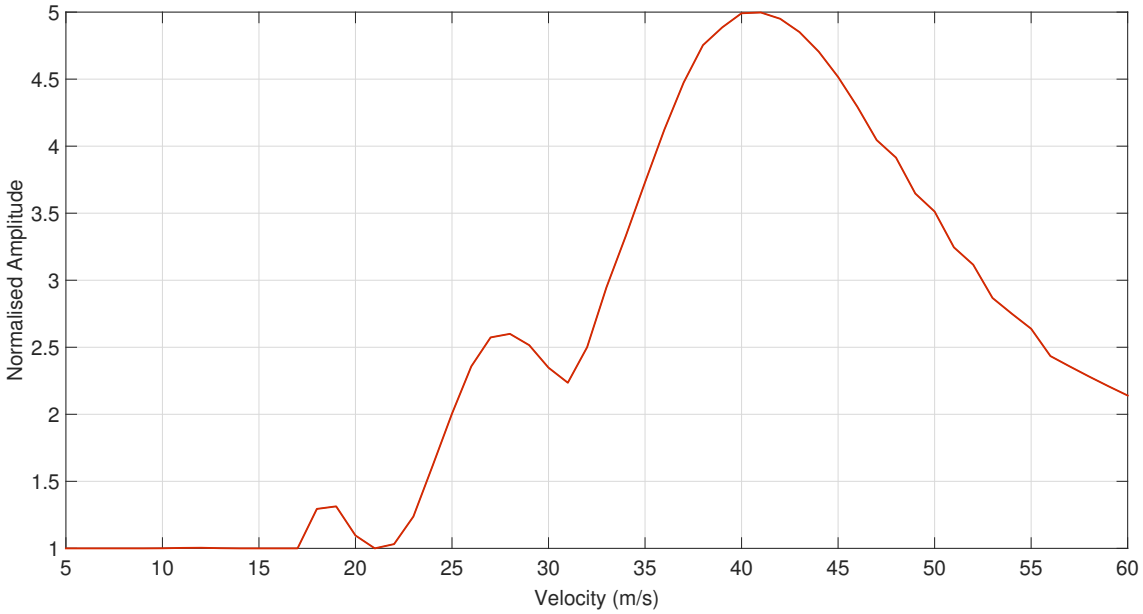


Figure 9.3.2: The IVS corresponding to the entire recording of length 1 second. The velocities were calculated assuming an interelectrode spacing of 7.5 mm and three distinct velocity peaks at 18, 27 and 41 m/s are visible, corresponding to the axonal diameters of 10, 15, and 22 μm respectively. The time domain SNR was 20 dB and the profile of the noise was AWGN.

populations with peak conduction velocities of 18, 27 and 41 m/s, corresponding to the simulated axonal diameters of 10, 15 and 22 μm respectively. The first electrode channel was chosen as the basis for classification using spike sorting, the raw data first being partitioned into individual APs using the NEO and time aligned to the point of maximal slope. The first two PCA scores were then computed and are plotted in Figure 9.3.3 as a scatter plot. k-Means clustering has been applied to partition the data into three distinct clusters. Each score has been colour coded to show the conduction velocities associated with each score as found using VSD: blue for 41 m/s, green for 27 m/s and red for 18 m/s. The PC scores clearly partition into three discrete clusters that correspond to the three axonal diameters present within the time recording. A similar result is obtained when the principal components are extracted from the remaining eight bipolar channels. Furthermore, in this case, it can be seen by inspection of Figure 9.3.3 that the classification process in fact only requires the first principal component score as there is very little variation along the second score. Clearly the spike sorting process has classified the APs into the same three partitions as VSR/VSD, but this important result is not altogether surprising as it is well-known that the diameter of the axon defines both the conduction velocity and the shape of the AP. Also visible in Figure 9.3.3 is an increase in the inter-score spread for the APs with a lower conduction velocity (and thus peak amplitude within the time domain).

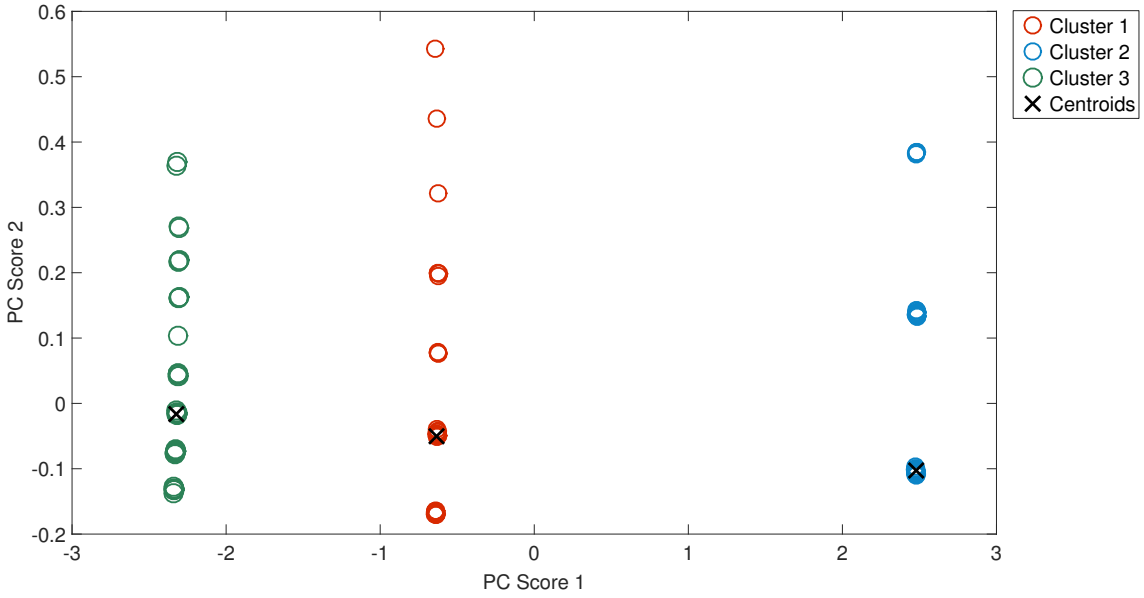


Figure 9.3.3: The first two PCA scores calculated from a single channel of simulated data containing 160 APs of three different axonal diameters. The scores partition very obviously into three classes corresponding to the velocities/fibre diameters identified by VSR in Figure 9.3.2. The three classes are clearly visible showing that only the first score is needed for complete classification of the data in this case. k-Means clustering has been used to identify each cluster and these have been colour coded, the centroids of each cluster are also shown.

9.3.2 Classification with Additive White Gaussian Noise

In order to compare the robustness of the two approaches under sub-optimal noise conditions the SNR was decreased from 30 dB to 1 dB using AWGN and the simulations were repeated in the same manner described previously. The IVS was computed for the new level of SNR and is shown in Figure 9.3.4 overlaid with the IVS corresponding to an SNR of 30 dB. The increased noise power results in a noticeable degradation of the IVS profile and a reduction in velocity selectivity for the three populations. However the three populations are still well defined and clearly visible, the effective boost in SNR provided by the delay-and-add process of $\sqrt{N}$ in this case provides an effective SNR of approximately 3 dB. The principal components were extracted and clustered using the same method as before, however in the case of spike sorting the same reduction in SNR has a very radical effect on the computed scores, as illustrated in Figure 9.3.5. In this case the individual scores are orders of magnitude different from their previous values and the k-means clustering algorithm has failed to correctly identify any distinct clusters. PCA based spike sorting appears, on this evidence, to have a significantly higher sensitivity to AWGN than a VSR-based approach. Note that extending the clustering process to higher dimensions (i.e considering more than the first two principal components) did not improve the accuracy of the spike sorting process.

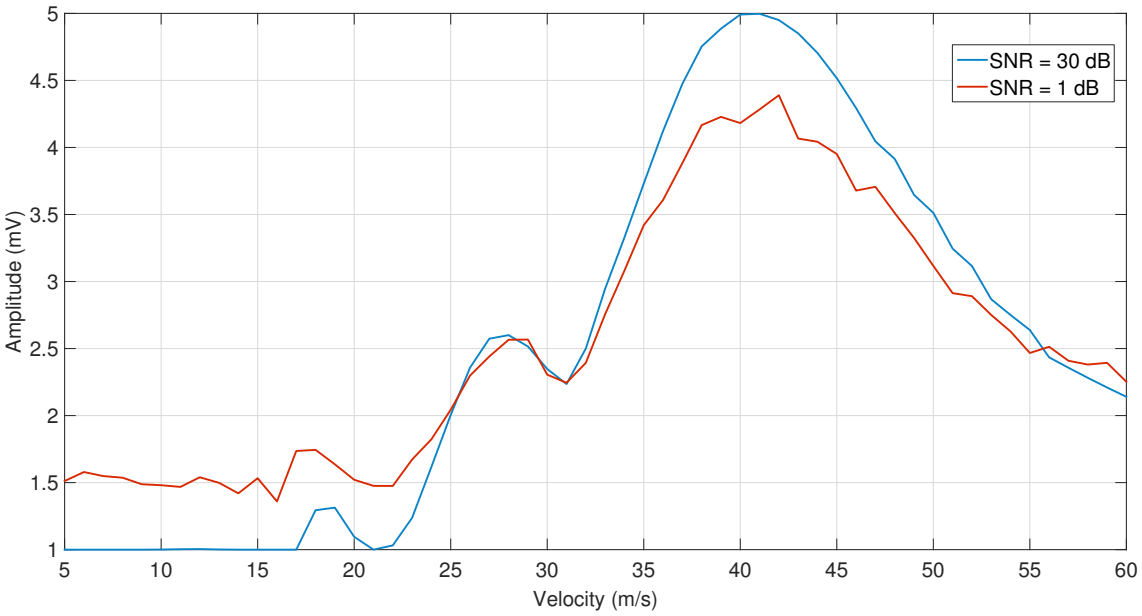


Figure 9.3.4: The effect of increased noise levels on the IVS, the SNR was reduced from 30 dB to 1 dB and the effect on the IVS is illustrated for both noise levels. Note that although the effect of the decrease in SNR is significant, the three component velocities are still clearly visible as distinct peaks within the IVS.

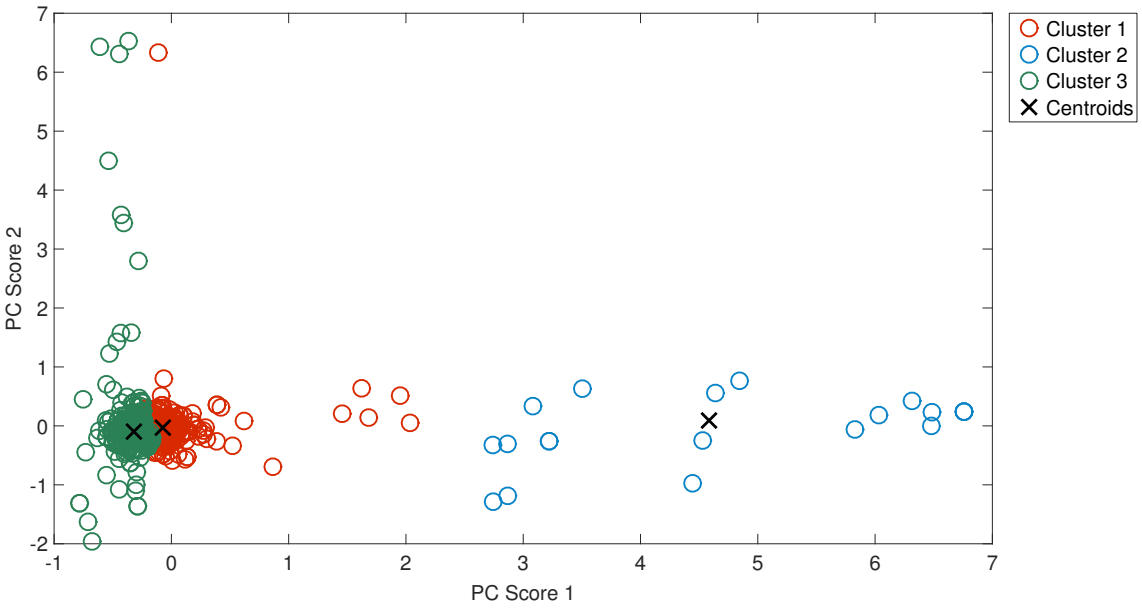


Figure 9.3.5: The effect of increased noise levels on PCA. As in Figure 9.3.4 the SNR was reduced from 30 dB to 1 dB and the effect on PCA investigated. Note that in contrast to the IVS in Figure 9.3.4, PCA scores 1 & 2 are now completely different in terms of scale. In particular, for an SNR of 1 dB, the scores are significantly more scattered and k-means clustering has failed to identify three unique clusters, in this case the use of PCA based spike sorting has failed.

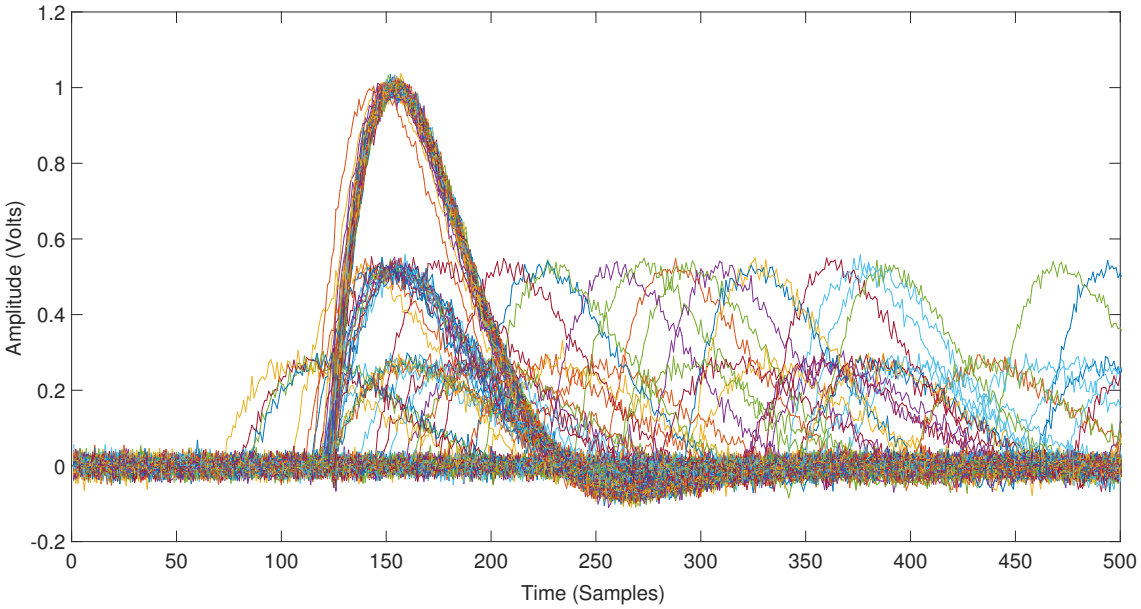


Figure 9.3.6: The effect of AWGN on the spike alignment process. Shown here are all 160 APs after extraction and alignment using the point of maximal slope, the large amplitude APs have been approximately aligned whereas the lower amplitude APs have been poorly aligned in time, a critical factor in PCA based analysis.

9.3.3 Classification with Ornstein-Uhlenbeck Process Noise

A step-by-step analysis of the spike sorting process for low levels of SNR reveals the largest source of error is introduced by the temporal alignment process, which as discussed in Chapter 4 is of critical importance with this type of feature extraction. In this study the point of maximal slope was used to align the AP waveforms, this method being one of the most commonly used for extracellular recordings [1]. The point of maximal slope was calculated using the centre difference operator and the presence of broad-band noise, such as AWGN, has the potential to significantly perturb the point of maximal slope. Visual examination of the alignment process as shown in Figure 9.3.6 demonstrates the problem; all 160 APs have been aligned using the point of maximal slope but APs with lower amplitudes are incorrectly aligned. For the purposes of illustration the data shown have an SNR of 15 dB.

The problems associated with temporal alignment of APs have been discussed in Chapter 4 and a novel method that attempts to solve many of these problems has been suggested in the use of the *centroid filter*. As the use of the centroid is not common practice in conventional spike sorting, and because the aim of this simulation study is to compare existing methods, the simulations were continued using the point of maximal slope as the alignment method. It was also discussed in Chapter 4 that uncorrelated broad-band noise sources such as AWGN might not be the most representative of biological recordings. So the type of noise used for the simulations was modified to follow an *Ornstein-Uhlenbeck* (OU) process as described by Equation 4.3.3. The time constant τ was 0.01 and the exper-

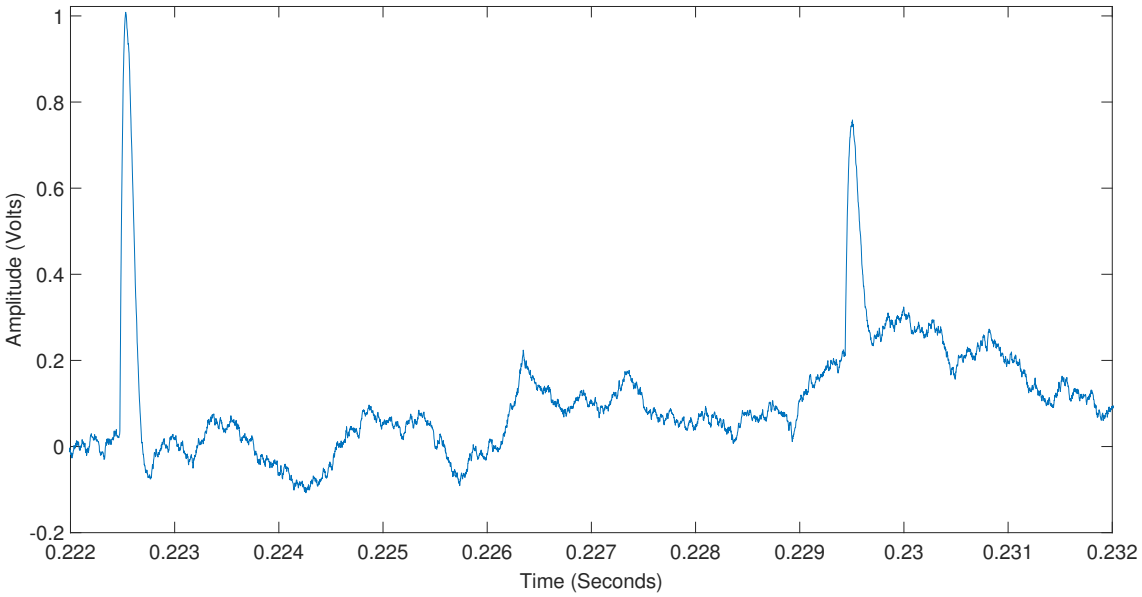


Figure 9.3.7: A snippet from a single dipole illustrating the presence of OU noise with an SNR of 1 dB, as found by computing the ratio of the summed squared signal magnitude to that of the noise. Unlike white noise the OU process is relatively band limited and displays characteristics typical with experimental data observed in recordings made from *rat*.

iments were repeated for an SNR of 1 dB. Figure 9.3.7 shows a snippet of a single channel illustrating the basic time domain noise profile for OU noise, the band limited nature of the OU noise process displays similar characteristics to data recorded experimentally from *rat* and detailed in Chapter 6. Figure 9.3.8 shows the IVS plots for both 30 dB and 1 dB SNR with OU noise and Figure 9.3.9 is the clustering result at an SNR of 1 dB. As with AWGN the reduction in the SNR has a noticable effect on the IVS, however the effect is substantially less than that for AWGN for the same SNR. The same is true for spike sorting, clearly the effect of the OU noise on the classification process is greatly reduced when compared to AWGN for the same levels of SNR. Figure 9.3.9 illustrates that the first two PCA scores are required for complete classification. The band limited nature of the OU noise allows for more robust temporal alignment than that observed with AWGN.

9.4 Results for Data Recorded *Ex Vivo* from *Frog*

In recent work, the use of VSR to discriminate between populations of *compound action potentials* (CAPs) in an explanted nerve from *frog* was demonstrated [3]. The data recorded by Schuettler et al. is used here with permission of the original authors. In this study polyimide thin-film technology was employed to produce two self-curling cuff electrodes. A 300 nm thick film of sputtered platinum was sandwiched between two 5 μm thick plates of polyimide. The two cuffs were 40 mm with an internal diameter of 1.5 mm, the platinum was patterned by photolithography in order to form the electrode contact, integrated tracks and bonding pads. A more detailed description of the fabrication process

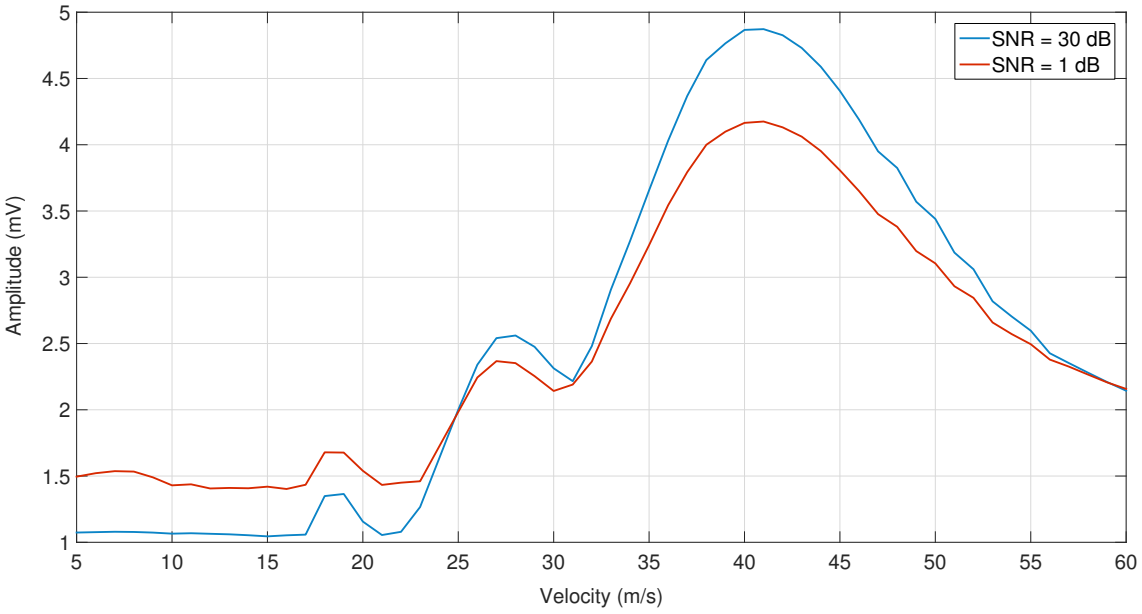


Figure 9.3.8: The effect of increased noise levels on the IVS. The SNR was reduced from 30 dB to 1 dB, this time using OU noise, the effect on the IVS is illustrated for both noise levels. Note that although the effect of the decrease in SNR is significant, the three component velocities are still clearly visible. The effect of the OU noise process is substantially less than that for AWGN for the same SNR (see Figure 9.3.4).

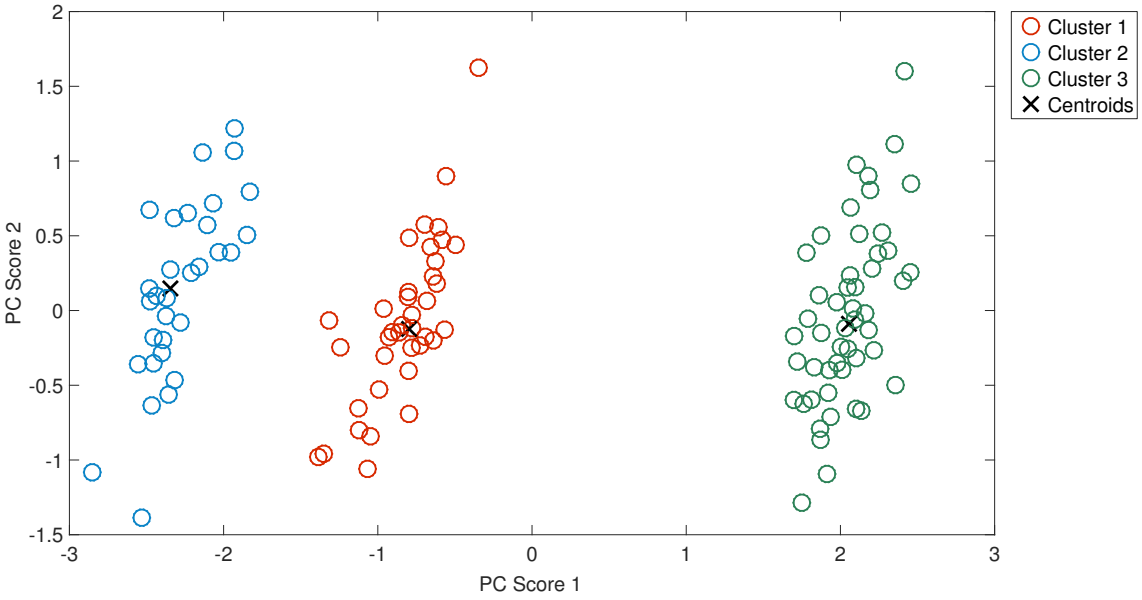


Figure 9.3.9: The effect of increased noise levels on PCA based spike sorting. As in Figure 9.3.5 the SNR was reduced from 30 dB to 1 dB but this time using an OU noise process. Note that in contrast to the result for AWGN given in Figure 9.3.5, PCA scores 1 & 2 now are accurately clustered intro three groups. The three classes are clearly visible showing that only the first two scores are needed for complete classification of the data in this case. k-Means clustering has been used to identify each cluster.

can be found in [7]. Each cuff was formed with eleven contact rings, each with a width of 500 μm and a centre-to-centre spacing of 3.5 mm. A further set of polyimide cuffs was used for electrical stimulation of the nerve. These cuffs were fabricated in the same way as the cuffs used for recording but were shorter, their total length was 12 mm, having three electrode rings of 0.2 mm width and a centre-to-centre pitch of 5 mm. Within the stimulation cuffs the individual electrodes were coated with platinum black in order to increase the charge injection capacity.

In the experiments performed by Schuettler et al. 80 to 90 mm long sections of the sciatic nerves were removed from decapitated *Xenopus Laevis frogs* and immersed in amphibian Ringer's solution [8] at room temperature. After removal of the connective tissue and small branches, the nerves varied in diameter from about 0.5 mm, close to the foot, to 2.0 mm, close to the spine. The nerves were stimulated electrically at the distal end so that CAPs propagated towards to the proximal end of the nerve, where the recording electrode was placed. Four sciatic nerves taken from three *frogs* were used. In the original experiment each nerve underwent electrical stimulation and *anodal blocking*. Anodal blocking is a method by which the usual law of inverse recruitment (that large, fast, axons are recruited first) may be circumvented. The nerve is stimulated at one site and the propagation of fast APs is blocked by an anodal pulse applied at a second location distal to the stimulation site. With a judicious choice of timing the fast APs can be blocked by the anodal pulse while the slow APs are allowed passed. Using this technique it was possible to generate predominantly a) high conduction velocity, b) high and low conduction velocity and c) low conduction velocity activity within the nerve. Since the main purpose of this chapter is to examine and compare the characteristics of two types of neural discriminators (VSR and spike sorting), the data sets employed were sets a) and b) whereby a high stimulating current was applied in conjunction with anodal blocking so as to recruit two different populations individually. The pulse width of the stimulator was set to 200 μs and the stimulation amplitude set to 2.2 mA. Sixty successive stimuli were applied in each condition and stimulation intensity.

Figure 9.4.1 shows each AP after extraction and time alignment using the NEO and the point of maximal slope respectively. VSD has been used to find the conduction velocity for each of the APs and they have been colour coded: red for 40 m/s and blue for 20 m/s. It is clear that the main variation within the shape of the two populations is in the *refractory* phase. Next, spike sorting using PCA was applied to a single electrode channel (proximal to the stimulating electrode) for the 124 APs shown in Figure 9.4.1. The first two PC scores were plotted for each AP in the scatter plot shown in Figure 9.4.2. Each AP has been colour coded to show the conduction velocity as determined by VSD: red for 40 m/s and green for 20 m/s. The important result is that (with the exception of the three outliers indicated) each AP has been grouped into one of two categories by both methods, with remarkable agreement. Spike sorting methods were also applied to the other eight channels and in almost all cases at least three principal components were required to separate the

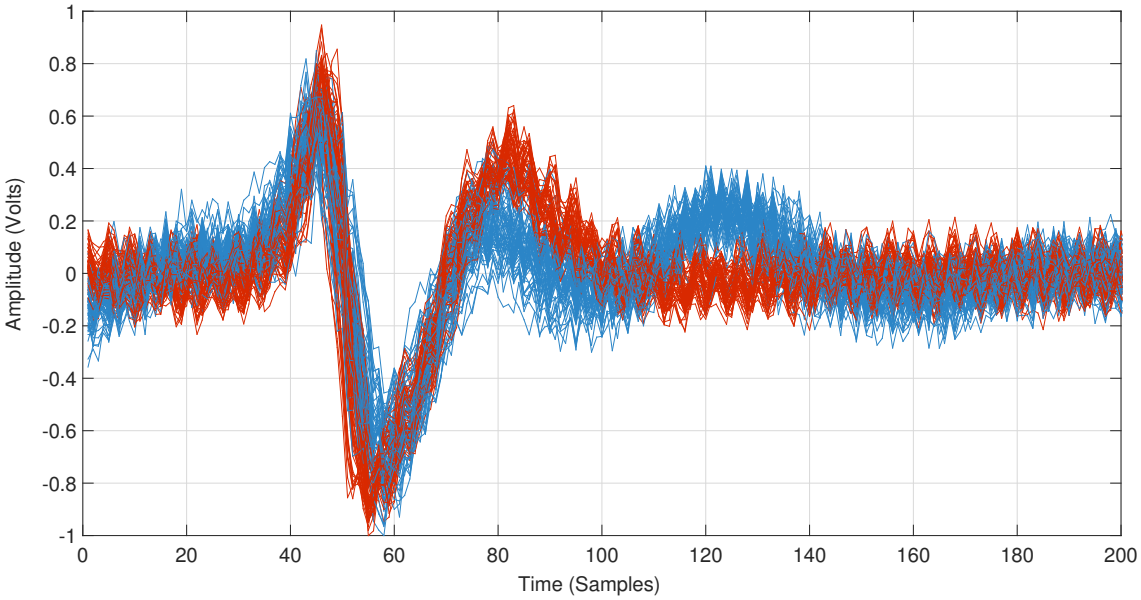


Figure 9.4.1: Extracted and aligned APs from a single channel of experimental data recorded from *frog*. The blue traces correspond to a conduction velocity of 20 m/s and the red traces to a conduction velocity of 40 m/s.

data. This is in agreement with the result presented in [5], and increases the complexity of the clustering algorithms required to classify each AP.

9.5 Applicability to Chronic Recordings

There are of course a number of key differences when considering the type of signal processing to apply to neural recordings. Perhaps the most distinct difference in requirement between a VSR based system and a traditional spike sorting system is the requirement for multiple electrodes; conventional spike sorting can operate on a single channel of data whereas VSR requires at least two (although more often five to ten) electrodes spaced equidistantly along a recording structure [1, 9]. Of course the increase in effective SNR obtained through the use of VSR with a multiple electrode system may enable recording from a nerve in a situation that is simply not possible with a single electrode system due to SNR constraints. Another key limitation of spike sorting methods is that they require the individual APs to be visible and nonoverlapping within the time domain recording, two constraints that limit the size of the fascicle that may be recorded from [10]. VSR has already been reliably demonstrated with electrically evoked overlapping APs [3, 11]. The use of clustering methods such as k-means requires the prior knowledge of how many different populations are expected, this level of knowledge would likely be patient specific. Whilst it has not been definitively quantified, the hardware implementation costs associated with morphological spike sorting almost certainly outweigh those associated with VSR. PCA requires complex matrix manipulations in order to compute singular value decompositions, an iterative and computationally expensive process. The delay-and-add

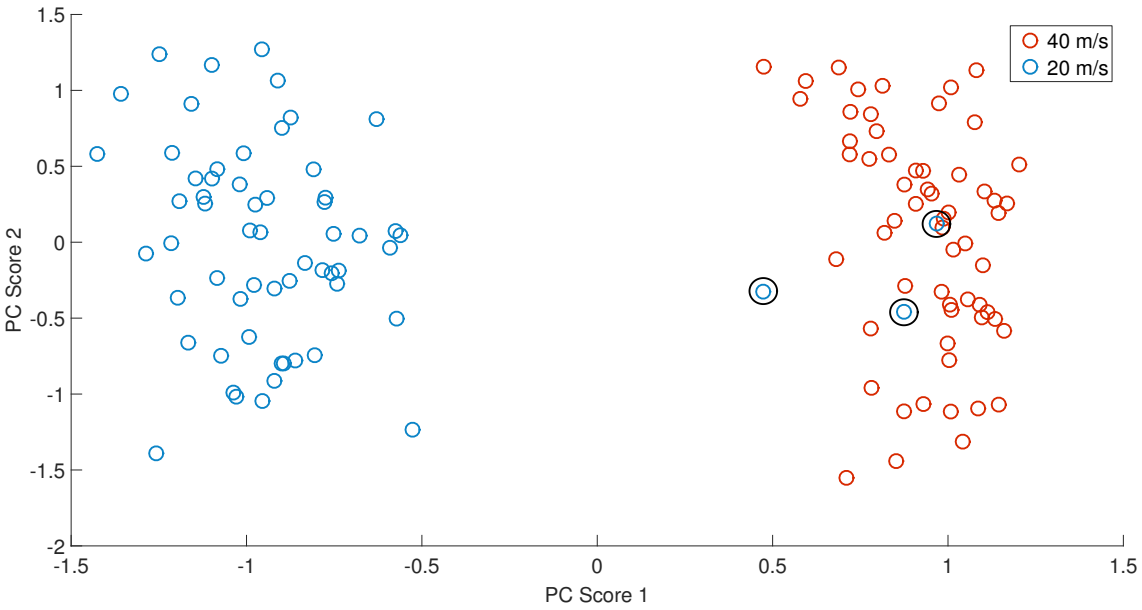


Figure 9.4.2: Scatter plot of the first two largest PCA scores for 124 APs recoded from *frog* using electrical stimulation with anodal blocking. The blue markers correspond to APs with conduction velocity 20 m/s and the red markers to APs with conduction velocity 40 m/s. Outliers (misidentified APs) are shown circled.

process required to form a simple VSR system by comparison is effectively real time and requires only the most basic signal processing operations.

Whilst it is unlikely that there would exist a recording scenario in which both spike sorting and VSR were applicable the results in this chapter have shown that both methods classify neurons based on axonal diameter and could feasibly be used interchangeably depending on electrode and signal processing constraints. The use of VSD would enable the creation of spike trains in very much the same way as spike sorting.

9.6 Conclusions

This chapter introduced a comparison between the methods of VSR and PCA-based spike sorting of neuronal recordings. A series of simulations of randomly distributed APs from a total of three distinct axons were performed with different levels of SNR and using both AWGN and OU process noise models, each having different effects on the performance of the two methods. For moderate SNRs ($>$ 30 dB AWGN, $>$ 10 dB OU) the VSR and PCA-based spike sorting methods classified APs identically, an important but unsurprising result as it is well-known that both the conduction velocity and the morphology of an AP is related directly to the diameter of the axon that produced it. When considering lower levels of SNR there is a distinct benefit from VSR based methods whereby the effective SNR is increased by a ratio of approximately $\sqrt{N}$, where N is the number of electrode channels [12]. The type of noise model applied had marginal effect on the VSR based methods, however this is not the case for spike sorting wherein the temporal alignment

process rapidly degenerated in the presence of broad-band noise such as AWGN. It has been suggested that the band-limited OU type noise models are more realistic of the background activity seen in biological recordings [5]. The use of OU type noise enables conventional PCA spike sorting methods to operate at significantly lower levels of SNR whilst retaining the simple alignment metrics such as the point of maximal slope. In Chapter 4 the centroid filter was proposed as a more reliable method for spike alignment and its use should enable the reduction in sensitivity of spike sorting methods to any kind of noise.

Data recorded using electrical stimulation and anodal blocking from an explanted nerve of *frog* was subjected to the same processing methods as the simulated data. It was shown that for 124 APs recorded from two distinct populations (axonal diameters) the PCA based spike sorting methods successfully classified all but 3 APs. The VSR based methods that were used to analyse the original data set successfully classify all of the APs into two distinct populations of 40 m/s and 20 m/s conduction velocity.

References

[1] S. Gibson, J. W. Judy, and D. Markovic, "Spike Sorting: The First Step in Decoding the Brain," *IEEE Signal Processing Magazine*, vol. 29, pp. 124–143, Jan. 2012.

[2] K. Famm, B. Litt, K. J. Tracey, E. S. Boyden, and M. Slaoui, "Drug discovery: a jump-start for electroceuticals.," *Nature*, vol. 496, pp. 159–61, Apr. 2013.

[3] M. Schuettler, N. Donaldson, V. Seetohul, and J. Taylor, "Fibre-selective recording from the peripheral nerves of frogs using a multi-electrode cuff.," *Journal of neural engineering*, vol. 10, June 2013.

[4] F. Murtagh and A. Heck, *Multivariate Data Analysis*, vol. 131 of *Astrophysics and Space Science Library*. Dordrecht: Springer Netherlands, 1987.

[5] D. A. Adamos, E. K. Kosmidis, and G. Theophilidis, "Performance evaluation of PCA-based spike sorting algorithms.," *Computer methods and programs in biomedicine*, vol. 91, pp. 232–44, Sept. 2008.

[6] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, pp. 129–137, Mar. 1982.

[7] T. Stieglitz, H. Beutel, M. Schuettler, and J. Meyer, "Micromachined, Polyimide-Based Devices for Flexible Neural Interfaces," *Biomedical Microdevices*, vol. 2, no. 4, pp. 283–294, 2000.

[8] A. Flores, M. Le, R. Vega, and E. Soto, "Histochemistryand role of nitric oxide synthase in the amphibian (Ambystoma tigrinum) inner ear," *Neuroscience Letters*, vol. 205, pp. 131–134, 1996.

[9] J. Taylor, M. Schuettler, C. Clarke, and N. Donaldson, "A summary of the theory of velocity selective neural recording.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2011, pp. 4649–52, Jan. 2011.

[10] K. H. Kim and S. J. Kim, "Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier.," *IEEE transactions on bio-medical engineering*, vol. 47, pp. 1406–11, Oct. 2000.

[11] M. Schuettler, V. Seetohul, J. Taylor, and N. Donaldson, "Velocity-selective recording from frog nerve using a multi-contact cuff electrode.," *Conference Proceedings: IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 2962–5, Jan. 2006.

[12] N. Donaldson, R. Rieger, M. Schuettler, and J. Taylor, "Noise and selectivity of velocity-selective multi-electrode nerve cuffs.," *Medical & biological engineering & computing*, vol. 46, pp. 1005–18, Oct. 2008.

Chapter 10

Conclusions and Future Directions

10.1 Conclusions

The main purpose of the work presented in this thesis was to extend the current state of the art of physiological *electroneurogram* (ENG) recordings using the method of *velocity selective recording* (VSR). If more advanced neural prostheses are to be manufactured, a stable and effective neural recording interface is required in order to provide closed loop feedback. VSR has been applied successfully to the analysis of electrically evoked *compound action potentials* (CAPs) in existing work and the work reported in this thesis attempts to extend these methods to the analysis of *physiological* (naturally occurring) ENG using chronically implantable electrode structures. The main questions that have been addressed are:

1. How can VSR be extended to extract the firing rate for particular neurons based on conduction velocity, which is an essential requirement for any neural recording system?
2. How can power efficient microelectronic signal processing methods be designed that are suitable for power critical applications in implantable neural interfaces?
3. What are the potential applications of VSR in the context of traditional morphological spike sorting and how do the two methods classify neuronal activity?
4. How can existing simulation systems be modified so that the design of amplifiers and signal processors for VSR systems can be improved without the use of expensive experimentation in animal models?

10.2 Significant Contributions of the Thesis

The following can be identified as significant contributions of this thesis to the advancement of neural signal recording:

1. Chapter 6 details recordings of the physiological ENG that were made *in vivo* from *rat* using multiple hook electrodes connected as isolated dipoles and immersed in oil. These recordings show individually identifiable *action potentials* that were analysed automatically using newly developed signal processing methods. Modulation of the ENG was elicited via direct stimulation of the *derma* of the animal and statistically significant changes in the recorded ENG were shown to encode the stimulation event. These recordings represent the first application of velocity selective recording to physiological ENG and demonstrate that encoded information, in this case skin sensation, can be reliably identified from neural recordings using minimal signal processing.

2. Pioneering recordings of both electrically evoked and physiological ENG were made from the intact right vagus nerve of *pig*. These recordings, a preliminary analysis of which is presented in Chapter 7, were made with a chronically implantable electrode structure, a cuff electrode, and extend the work presented in *rat* to data obtained with an implantable electrode structure. Although only a preliminary analysis of these recordings is presented the results show that it is possible to observe APs even in a large nerve, such as the vagus nerve, with a *multiple electrode cuff* (MEC).

3. In Chapter 5 a new method was described that enables the firing rates of individual neurons to be extracted in real time using a velocity based signal processor. Previously available methods for analysing the velocity content of neural recordings provide no method by which firing rates could be extracted, the new method that has been developed extends the current methods to provide this functionality and has been validated using physiological recordings of the ENG made in *rat*. The extraction of firing rates is key to the development of a neural recording interface.

4. A novel discrete time filter has been developed in Chapter 4 that provides an efficient method for finding the *centre of gravity* of pulses. This filter was demonstrated as a method for the temporal alignment of APs and is substantially more resilient to both correlated and uncorrelated noise than existing methods. A multiplier free implementation has been described and detailed power and resource measurements were presented for a typical implementation. This filter has applications in both conventional spike sorting and the newly developed methods in VSR.

5. The methods of VSR are validated and contrasted in the context of traditional morphological spike sorting in Chapter 9. Both methods were applied to simulated data as well as to data recorded *ex vivo* from *frog* and it was shown that both methods classify neural activity based on axonal diameter and may be equally well

applied to the same data sets. The data recorded from *frog* were recorded in another study but were used here for comparative purposes.

6. In Chapter 8 a high speed hardware based simulation system was developed that extends existing technology using modern microelectronics systems to provide a high level of reconfigurability. The simulation system makes use of recent recordings made in rat to produce realistic test patterns for the design of velocity based signal processors. The simulator addresses a number of key issues with existing designs and aims to reduce the reliance on testing with animal models.

10.3 Future Directions and Challenges

There are many more challenges that must be overcome before a viable neural recording interface based on VSR can be constructed and implanted. The data presented in Chapter 7 from *pig* represent perhaps the most exciting result as it suggests that it may be possible to apply VSR to a large nerve using an MEC. However the analysis of this data is incomplete and as yet no definitive conclusions have been drawn. A detailed analysis is required that would reveal if it is possible to extract cardiac function or blood pressure from these physiological recordings. All the experimental work in the thesis was performed as acute experiments lasting no more than a few hours. Once the required electrode configuration for a particular target nerve has been described then long term chronic experiments are required in order to examine how the nerve-electrode interface varies over time with normal cellular functions, as well as to examine the mechanical stability of the electrode structures.

VSR has been applied to the analysis of physiological ENG that is predominantly sensory (afferent), and although theoretically feasible has yet to be applied to the analysis of data that is simultaneously afferent and efferent, such as would be expected in the vagus nerve. The signal processing tasks were performed offline using MATLAB but once the processing chain has been finalised for a particular application then these tasks need to be integrated onto an implantable *integrated circuit* (IC). An analysis needs to be performed in order to examine the relative merits of integrating the analogue amplifiers and the digital signal processing on a single (mixed signal) IC.

There may be power savings available in a mixed signal integration, as there is no requirement for an electrical interface between two chips, as would be the case for separate analogue and digital integration. However the design process would have to carefully control the noise aspects of a mixed signal design, especially in the presence of the very low noise high sensitivity front end amplifiers. Some of the digital processing may be done in the analogue domain, using switched capacitor delay lines for example, and this may further reduce power consumption.

Techniques such as clock gating could be used to control device activity, so that detailed

recordings and analysis only takes place at specified intervals, or after the detection of a trigger event. The level of processing performed on-chip will depend heavily on the efficiency of the transcutaneous communications system that is employed, and there inevitably will be a break even point at which data is best processed off-chip. Whilst power consumption may be reduced by using a modern deep-submicron process this may come at the cost of reduced device longevity, and these issues need to be further investigated. Device security will also play a key role in modern *bioelectronic* systems and devices that are attached to the nervous system will have to be exceptionally secure to outside or malicious interference, devices fitted with electrical stimulators for example may have the potential to inflict lethal damage if used incorrectly.

10.4 Concluding Remarks

Neural interfaces that are directly implanted within the peripheral nervous system of *man* represent an exciting and promising field with applications ranging from diagnostics and therapeutics to advanced human-machine interfaces. The nervous system is still widely regarded as one of the last frontiers of science and the potential applications of the so called bioelectronic medicines are substantial. The work undertaken in this thesis has extended the concepts of neural recording using VSR towards a chronically viable neural recording system and highlights the remaining steps that must be undertaken if this ultimate goal is to be achieved.